

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<p>(51) International Patent Classification ⁶ : G01N 33/574, C12Q 1/48, A61K 31/55, 31/275, 31/35, 39/395, 38/00, 31/70, G01N 33/573, C12Q 1/68, C12N 9/12, 15/11</p>	<p>A1</p>	<p>(11) International Publication Number: WO 95/02 (43) International Publication Date: 19 January 1995 (19.01.95)</p>
<p>(21) International Application Number: PCT/GB94/01480 (22) International Filing Date: 8 July 1994 (08.07.94) (30) Priority Data: 9314271.9 9 July 1993 (09.07.93) GB (71) Applicants (for all designated States except US): THE INSTITUTE OF CANCER RESEARCH [GB/GB]; 17A Onslow Gardens, London SW7 3AL (GB). THE WELLCOME FOUNDATION LIMITED [GB/GB]; Unicorn House, 160 Euston Road, London NW1 2BP (GB). (72) Inventors; and (75) Inventors/Applicants (for US only): GUSTERSON, Barry, Austin [GB/GB]; Institute of Cancer Research, Royal Cancer Hospital, Haddow Laboratories, 15 Cotswold Road, Sutton, Surrey SM2 5NG (GB). CROMPTON, Mark, Roger [GB/GB]; Institute of Cancer Research, Royal Cancer Hospital, Haddow Laboratories, 15 Cotswold Road, Sutton, Surrey SM2 5NG (GB). MITCHELL, Philip, John [GB/GB]; Institute of Cancer Research, Royal Cancer Hospital, Haddow Laboratories, 15 Cotswold Road, Sutton, Surrey SM2 5NG (GB). BARKER, Karen, Tracey [GB/GB]; Institute of Cancer Research, Royal Cancer Hospital, Haddow Laboratories, 15 Cotswold Road, Sutton, Surrey SM2 5NG (GB).</p>	<p>tories, 15 Cotswold Road, Sutton, Surrey SM2 5NG (GB). MARTINDALE, Jane, Elizabeth [GB/GB]; Institute of Cancer Research, Royal Cancer Hospital, Haddow Laboratories, 15 Cotswold Road, Sutton, Surrey SM2 5NG (GB). P. Martin, John [GB/GB]; The Wellcome Research Laboratories, Langley Court, South Eden Park Road, Beckenham, Kent BR3 3BS (GB). SPENCE, Paul [GB/GB]; The Wellcome Research Laboratories, Langley Court, South Eden Park Road, Beckenham, Kent BR3 3BS (GB). (74) Agent: MARCHANT, James, Ian; Elkington and Fife, Prospect House, 8 Pembroke Road, Sevenoaks, Kent TN13 3AB (GB). (81) Designated States: AM, AT, AU, BB, BG, BR, BY, CA, CN, CZ, DE, DK, ES, FI, GB, GE, HU, JP, KE, KG, KP, KR, KZ, LK, LT, LU, LV, MD, MG, MN, MW, NL, NO, NZ, PL, PT, RO, RU, SD, SE, SI, SK, TJ, TT, UA, UZ, VN, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).</p> <p>Published With international search report. Before the expiration of the time limit for amendment of claims and to be republished in the event of the record being amended.</p>	
<p>(54) Title: CELL GROWTH FACTOR RECEPTORS (57) Abstract The invention relates to a method of screening a substance for potential utility as a therapeutic agent in the treatment of cancer comprising providing a standard system in which a protein tyrosine kinase or an active fragment thereof is able to develop a measurable effect, allowing the protein tyrosine kinase to develop that effect in the presence and absence of the said substance and measuring the effect, ability to produce significant inhibition of the effect being taken as an indication of potential utility as a therapeutic agent, where the protein tyrosine kinase is characterised by the amino acid sequence of SEQ ID NO. 1 or an amino acid sequence showing a significant degree of homology thereto. The invention also relates to a therapeutic agent thus identified, to a method for detecting the protein tyrosine kinase and an enzyme substrate complex comprising the protein tyrosine kinase and a therapeutic agent capable of modulating the activity of the protein tyrosine kinase.</p>		

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	GB	United Kingdom	MR	Mauritania
AU	Australia	GE	Georgia	MW	Malawi
BB	Barbados	GN	Guinea	NE	Niger
BE	Belgium	GR	Greece	NL	Netherlands
BF	Burkina Faso	HU	Hungary	NO	Norway
BG	Bulgaria	IE	Ireland	NZ	New Zealand
BJ	Benin	IT	Italy	PL	Poland
BR	Brazil	JP	Japan	PT	Portugal
BY	Belarus	KE	Kenya	RO	Romania
CA	Canada	KG	Kyrgyzstan	RU	Russian Federation
CF	Central African Republic	KP	Democratic People's Republic of Korea	SD	Sudan
CG	Congo	KR	Republic of Korea	SE	Sweden
CH	Switzerland	KZ	Kazakhstan	SI	Slovenia
CI	Côte d'Ivoire	LI	Liechtenstein	SK	Slovakia
CM	Cameroon	LK	Sri Lanka	SN	Senegal
CN	China	LU	Luxembourg	TD	Chad
CS	Czechoslovakia	LV	Latvia	TG	Togo
CZ	Czech Republic	MC	Monaco	TJ	Tajikistan
DE	Germany	MD	Republic of Moldova	TT	Trinidad and Tobago
DK	Denmark	MG	Madagascar	UA	Ukraine
ES	Spain	ML	Mali	US	United States of America
FI	Finland	MN	Mongolia	UZ	Uzbekistan
FR	France			VN	Viet Nam
GA	Gabon				

CELL GROWTH FACTOR RECEPTORS

5 The present invention relates to cell growth factor receptors. More particularly the invention relates to the use of a tyrosine kinase growth factor receptor in the development of diagnostic and therapeutic approaches to cancer, for example breast cancer.

10 Protein tyrosine kinases are enzymes which show the property of catalysing the transfer of phosphate groups from donor molecules (ATP) to the hydroxyl groups of tyrosine residues in polypeptides. Known tyrosine kinases can be
15 classified into two broad groups. Transmembrane tyrosine kinases traverse cellular membranes so that they possess extracellular and intracellular domains. Cytoplasmic tyrosine kinases are located only intracellularly. A general feature of transmembrane (also referred to as receptor) tyrosine
20 kinases is that they possess extracellular ligand-binding domains, hydrophobic transmembrane sequences, and intracellular portions which include the tyrosine kinase domains (for a review see Ullrich & Schlessinger Cell, 61, 203-212 (1990)).

25 Tyrosine kinases can induce cell proliferation, cell transformation and regulate developmental events (see reviews by Hanks et al, Science, 241, 42-75 (1988) and Cantley et al, Cell 64, 281-302 (1991)). In general, where functional assays have been available (e.g. mitogenesis or transformation), it
30 has been shown that the biological functions of tyrosine kinases are usually dependent on intact enzyme activity and that through autophosphorylation and the phosphorylation of other proteins, they alter the subcellular localisation and activities of various components of the intracellular
35 signalling pathways.

Analysis of the oncogenes of many acutely transforming animal retroviruses has revealed that their products frequently manifest tyrosine kinase activity, as do the

products of their cellular proto-oncogene counterparts. Other genes encoding tyrosine kinases have been found to be altered by DNA rearrangements in cancer cells, the result being a presumed acquisition of cellular transforming activity (for example c-abl, c-met). Some cellular proto-oncogenes encoding tyrosine kinases have been cloned independently by virtue of the fact that they encode growth factor receptors, for example the epidermal growth factor receptor. On the other hand, the genes for other growth factor receptors which have tyrosine kinase activities, such as those for platelet-derived growth factor and insulin-like growth factors, have been well characterised, but have never been found to be transduced by retroviruses.

The c-erbB-2/HER2/c-neu gene encodes a transmembrane receptor-like tyrosine kinase which is structurally very similar to the EGF/TGF alpha receptor. The c-erbB-2 gene has been found to be overexpressed in 20 to 30% of human breast tumours, often in association with gene amplification, and this phenotype is now generally accepted as predictive of poor disease free and overall survival (for reviews see Sunderland & McGuire in *Regulatory Mechanisms in Breast Cancer*, Lippman & Dickson (Eds.), Kluwer Academic Publishers, Boston, pages 3 to 22 (1991) and Gusterson et al., *J. Clin. Oncol.*, 10, 1049-1056 (1992)). Overexpression of the gene in fibroblasts induces transformation (di Fiore et al, *Science*, 237, 178-182 (1987) and Hudziak et al, *Proc. Natl. Acad. Sci. (USA)*, 84, 7159-7162 (1987)), and ligands which bind to the c-erbB-2 gene product (human and rodent) and activate its tyrosine kinase activity have been identified (Lupu et al, *Science*, 249, 1552-1555 (1990), Dobashi et al, *Proc. Natl. Acad. Sci. (USA)*, 88, 8582-8586 (1991), Wen et al, *Cell*, 69, 559-572 (1992) and Holmes et al, *Science* 256, 1205-1210 (1992)). Some of these ligands increase the rate of proliferation of cells expressing the c-erbB-2 protein. The strong implication of these and other findings is that overexpression of this tyrosine kinase in some breast tumour cells is one of the important steps in their progression towards tumourigenicity, and therefore that c-erbB-2 can function as an oncogene.

A variety of other receptor tyrosine kinases, including those for the IGFs and the FGFs, are expressed in breast tumours (Stewart et al, J. Biol. Chem., 265 21172-21178 (1990) and Wellstein & Lippman in Molecular Foundations of Oncology, Broder (Ed.), Williams and Wilkins, Baltimore, pages 403-418 (1991)), and the EGF/TGF alpha receptor is overexpressed in some cases (for example Horak et al, Oncogene, 6, 2277-2284 (1991)). As with c-erbB-2, overexpression of EGFs has been correlated with poor prognosis. Since the ligands of some of these receptors can be detected in breast tumour samples, it may be that breast tumour development is also regulated by these molecules (for a review see van de Vijver & Nusse, Biochim, Biophys. Acta, 1072, 35-50 (1991)).

The features which are ideally required in a molecular target for tumour therapies include preferential expression in the tumours, a role in the development of the tumours (rather than merely being a marker), and a knowledge of the mechanism of its action as a starting point for the rational design of activity modulators. It appears that all of these three criteria are satisfied, at least to some extent, in the case of c-erbB-2 and breast cancer and a considerable research effort is currently being directed to this molecule.

It has recently been found that the alkaloid K252a, at appropriate concentrations, selectively inhibits the tyrosine kinase and biological activities of the trk-class of neurotrophin receptors, but has no effect on the activities of v-src, v-fms or the receptors for EGF and PDGF, and also does not affect general cell viability or proliferation (Tapley et al, Oncogene, 7, 371-381 (1992)). Members of another class of molecules, the tyrphostins, may be selective inhibitors of the EGFr/c-erbB-2 class of tyrosine kinases (Gazit et al, J. Med. Chem., 34, 1896-1907 (1991)). An antibody which inhibits the proliferation of human tumour cells by binding to the c-erbB-2 protein is currently under evaluation as a potential cancer therapy (Carter et al, Proc. Natl. Acad. Sci. (USA), 89, 4285-4289 (1992)).

If the transformation of a significant minority of breast tumour cells is due, at least in part, to the disregulated activity of the c-erbB-2 tyrosine kinase, then it seems likely that other breast tumours owe aspects of their transformed phenotype to the activity of other tyrosine kinases. The identification of further tumour associated tyrosine kinases would allow the development of novel diagnostic and therapeutic approaches to cancer and in particular breast tumours.

The cDNA cloning of a previously novel receptor tyrosine kinase from breast carcinoma cells has recently been reported (Johnson et al., Proc. Natl. Acad. Sci. USA, 90, 5679-5681 (June 1993)) and the DNA and predicted amino acid sequence of this protein are set out in

SEQ ID NO. 1.

The predicted amino acid sequence is set out separately in

SEQ ID NO. 2.

The authors designated this protein tyrosine kinase DDR (Discoidin domain receptor) and detected the transcription product of the DNA sequence in other human breast carcinoma cell lines and the translation product (protein) was also detected in breast carcinoma cell lines. However, the authors do not suggest that the protein tyrosine kinase plays any role as a tumour causative agent and no potential utility is suggested in the therapy or diagnosis of cancer, in particular human breast cancer.

Another recent disclosure suggests that the same protein tyrosine kinase has been detected in an epithelial ovarian cancer cell line (Laval et al, American Association for Cancer Research, Meeting, May 1993, Abstract No. 3144).

It has now been found that the receptor tyrosine kinase referred to above is expressed at a much higher level in human breast tumour cells than in normal breast. This suggests an

involvement of the tyrosine kinase in the transformation of breast cells. The present invention relates to the use of the tyrosine kinase as defined above in the diagnosis and/or therapy of cancer, particularly human breast tumours.

5

The present invention involves the use of a DNA isolate encoding a protein tyrosine kinase having the amino acid sequence shown in

SEQ ID NO. 1

10 or an amino acid sequence showing a significant degree of homology thereto, preferably at least 60% homology in the catalytic domain and at least 40% homology in the remainder of the molecule, more preferably at least 90% homology overall, for example at least 95% homology, or a fragment
15 thereof.

According to one embodiment, the DNA isolate encodes all of the amino acid sequence shown in the SEQ ID NO. 1 or an amino acid sequence showing a significant degree of homology
20 thereto. According to another embodiment the DNA isolate encodes a fragment of the amino acid sequence of the protein tyrosine kinase or an amino acid sequence showing a significant degree of homology thereto. DNA sequence encoding fragments of the protein tyrosine kinase preferably encodes
25 those parts of the amino acid sequence which characterise the enzyme, i.e. those parts which are most distinct from other protein tyrosine kinases. Most preferably the DNA sequence encodes all or part of the catalytic domain of the enzyme, i.e. amino acid residues 617 to 907 as shown in SEQ ID NO. 1.

30

The DNA isolate may have the base sequence defined in SEQ ID NO. 1 to encode the relevant amino acid sequence. Alternatively the DNA isolate may have any other DNA sequence encoding the relevant amino acid sequence.

35

The DNA isolate may take the form of a cloning vector or an expression vector preferably a plasmid vector, including DNA as defined above. In the case of an expression vector the DNA will be under control of an appropriate promoter and will

include regulatory elements required for expression in a suitable host cell.

5 The present invention also involves the use of a recombinant cell line transformed with an expression vector as defined above and capable of expressing a recombinant protein tyrosine kinase having amino acid sequence shown in

SEQ ID NO. 1

10 or an amino acid sequence showing a significant degree of homology thereto or a fragment thereof.

15 Suitable host cells include mammalian cells, insect cells, yeast cells and bacterial cells and the expression vector will include a promoter and other regulatory elements appropriate to the host cell in question. Preferred host cells include CHO cells, myeloma cells, primary immortalised human breast cells, rodent fibroblast cell lines, baculovirus cells, yeast and *E. coli* cells. General techniques for manipulating DNA coding sequences and expressing such

20 sequences in various types of cells are well known to those skilled in the art and are described for example in Sambrook et al, Molecular Cloning: A Laboratory Manual, 2nd Edition, Cold Spring Harbor Laboratory.

25 The present invention also involves the use of a protein tyrosine kinase having the amino acid sequence shown in

SEQ ID NO. 1

30 or an amino acid sequence showing a significant degree of homology thereto or a fragment thereof.

The protein tyrosine kinase or a fragment thereof may be a recombinant polypeptide produced as the expression product of a coding sequence as defined above in a recombinant cell line. Alternatively the protein tyrosine kinase may be

35 produced in a human tumour cell line. In either case the protein may be extracted and purified by standard techniques, for example antibody affinity chromatography. The glycosylation of the protein (if any) will depend on the cells in which it is produced and protein produced in human tumour

cells in culture should have glycosylation equivalent to that produced *in vivo* in human tumours.

5 Part of the DNA sequence encoding the protein tyrosine kinase set out in SEQ ID NO. 1 was isolated in the manner described in more detail in the example set out below but which involves briefly the following steps:

- i) isolation of mRNA from tumour metastatic tissue;
- ii) preparation of cDNA from the isolated mRNA;
- 10 iii) PCR amplification using degenerate oligonucleotide primers designed to amplify sequence associated with protein tyrosine kinases;
- iv) subcloning PCR products;
- v) identification of protein tyrosine kinase products
- 15 amplified at higher levels from tumour bearing tissue;
- vi) DNA sequencing and identification of novel protein tyrosine kinase sequence;
- vii) isolation of cDNA sequence 3' of the region amplified by
- 20 3' RACE PCR.

20 Application of the above protocol identified part of the DNA sequence for the protein tyrosine kinase which was designated PTK22 and which corresponds to the protein tyrosine kinase designated DDR by Johnson et al (Supra). As already

25 noted, the enzyme has now been shown to be overexpressed in certain human breast tumours as compared to normal breast tissue.

30 The present invention thus relates to the application of the tyrosine kinases for the development of therapeutic, prognostic and diagnostic approaches to cancer. The invention is particularly applicable to breast cancer, however, the same approaches may also be applicable to other cancers and research data suggests a connection between breast cancer and

35 ovarian cancer and also between these two cancers and gastric cancer.

In terms of therapy, the involvement of the tyrosine kinase in tumours means that beneficial clinical effects in

the treatment of tumours can be obtained by modulating the tyrosine kinase activity and/or the functionality of the receptor. There are a number of ways in which such modulation could be achieved.

5

The activity of the protein tyrosine kinase could be significantly impaired or inhibited by small chemical molecules and screens to identify suitable small molecular weight inhibitors can be developed as described in more detail
10 below. This approach is analogous to that proposed for other tyrosine kinases (see "Drugs of the Future", 17(2), 119-131, (1992)). A further embodiment of this approach would be to derive nmr spectra or crystal structures for the protein tyrosine kinases or domains thereof and use the structural
15 information so obtained to synthesise chemical structures de novo which could be similarly screened for activity as inhibitors.

The approach described above could be extended to
20 encompass small peptides which are either competitive for signalling action of the tyrosine kinase receptor, or which demonstrate useful binding thereto, such that the receptor function is inhibited. This would include, for example, peptides which are capable of blocking substrate binding to
25 the receptor, including but not limited to those which bear SH2 domains. An additional approach is to use small molecules or peptides to block or interfere with ligand activation or dimerisation of the tyrosine kinase. Both of these events are generally required to effect a biological signal so that
30 blocking or inhibition thereof could be used to therapeutic effect.

Macromolecules, such as antibodies, could also be used as inhibitory molecules to block receptor function. This
35 would be particularly applicable to antibodies raised against the extracellular domain of the protein tyrosine kinase where antibodies could promote internalisation or interfere with ligand binding. Either of these situations could produce significant modulation (positive or negative) of the activity

of the protein.

Protein tyrosine kinase function could also be blocked by expressing or administering a peptide which is a truncated or altered version of the protein. These are typically referred to as dominant negative proteins and are believed to sequester the active protein in a non-functional complex. Protein tyrosine kinase function could also be modulated by blocking translation of the mRNA encoding the protein using antisense oligonucleotides.

A further regulatory role of the protein tyrosine kinase could be manipulated by controlling expression or activity of specific phosphatases which control the degree of protein phosphorylation. It is recognised that the phosphorylation of specific residues, particularly tyrosine, threonine and serine residues, play an important role in transmitting a biological signal from the tyrosine kinase.

Use of the above therapeutic approaches depends on the identification and/or development of suitable agents for modulating, for example inhibiting, particular functions or activities of the protein tyrosine kinase. The present invention thus relates to assays or screens by which such agents can be identified.

The present invention provides a method of screening a substance for potential utility as a therapeutic agent in the treatment of cancer, in particular breast cancer, which comprises providing a standard system in which a protein tyrosine kinase or an active fragment thereof is able to develop a measurable effect, allowing the protein tyrosine kinase to develop that effect in the presence and absence of the said substance and measuring that effect, ability to produce significant inhibition of the effect being taken as an indication of potential utility as a therapeutic agent, wherein the protein tyrosine kinase is characterised by the amino acid sequence shown in

SEQ ID NO. 1

or an amino acid sequence showing a significant degree of homology thereto.

In accordance with this general principle, screens for potential therapeutic agents can take a number of forms. For example an initial screen to determine whether or not a substance merits further investigation as a potential inhibitor of the protein tyrosine kinase will usually be biochemical and should preferably be simple, rapid and capable of high through-put. Such screens will often make use of the protein expressed in a recombinant expression system or derived from a cellular source in which the protein is overexpressed. Particularly suitable recombinant expression systems include insect baculovirus, and the use of primary immortalised human breast cell lines, or rodent fibroblast cell lines as host cells. The protein is then used directly in an *in vitro* assay with and without the potential inhibitor. This gives rapid data on the performance of the substance as an inhibitor often providing direct evidence that the substance can inhibit the enzymic activity of the protein tyrosine kinase and thus merits further study. This type of isolated system can also provide data on specificity but will not generally provide any information on bioavailability. The enzymic activity of the protein tyrosine kinase in vitro can be measured, for example, by measuring autophosphorylation or phosphorylation of a model substrate, by following incorporation of radioactive phosphate or by using anti-phosphotyrosine antibodies.

Further information concerning potential inhibitors can be provided by cell based screens which make use of a phenotypic alteration, e.g. a change in morphology and/or tumorigenicity, conferred by expression of the protein tyrosine kinase in a recombinant cell line or overexpression in any other available cell line. Particularly suitable recombinant cell lines again use primary immortalised human breast cell lines or rodent fibroblast cell lines as host cells. The cells can be used in proliferation or tumorigenicity assays with and without the potential inhibitor

looking for substances able to halt growth of and/or morphologically de-transform the cells. A control for substances which are generally cytotoxic can be provided by the same parental cell line engineered to express another oncogene which transforms the cells at a signalling point downstream of the protein tyrosine kinase. Assays of this sort can be very informative and they may provide data on mode of action of inhibitors.

The final stage of screening is the development of animal screens. These are time consuming and expensive and for these and ethical reasons their use is kept to a minimum, however they can give vital information regarding metabolism, clearance and performance of a substance in a true *in vivo* system which cannot be obtained in other ways. Animal screens ideally use the same recombinant or other cell lines as described above which express the protein tyrosine kinase. The cells are used as xenografts in animals, for example, nude mice, to give tumours whose growth is controlled by expression of the protein tyrosine kinase. The mice are then given inhibitors to determine whether they are capable of causing the selective regression of the xenografts without blocking the growth of a control xenograft. Alternatively transgenic animals which develop tumours by virtue of the appropriate tissue specific expression of the target protein tyrosine kinase can be used in a similar manner.

The invention also extends to therapeutic agents identified by use of any or all of the screens referred to above. Preferably the therapeutic agent is a chemical molecule of relatively low molecular weight, for example, less than about 1000. Examples of suitable classes of molecule include staurosporine analogues, tyrphostins and flavonoids. Alternatively, the therapeutic agent can be a macromolecule, for example an antibody raised against the extracellular domain of the protein tyrosine kinase. The therapeutic agent can also be a peptide or an antisense oligonucleotide capable of blocking translation of the mRNA encoding the protein tyrosine kinase.

The invention also extends to an enzyme-substrate complex which comprises a protein tyrosine kinase characterised by the amino acid sequence shown in

SEQ ID NO. 1

5 or an amino acid sequence showing a significant degree of homology thereto and a therapeutic agent capable of modulating the activity of the said protein tyrosine kinase.

The protein tyrosine kinases can also be used in
10 prognostic and diagnostic applications. Thus antibodies raised against the protein tyrosine kinase or nucleic acid encoding the tyrosine kinase can be used as the basis for screening tissue, in particular tumour tissue, for the presence of the protein tyrosine kinases. If, as in the case
15 of c-erbB-2, the expression of the tyrosine kinase correlates with poor prognosis, then treatment of the patient could be modified accordingly. Similarly, again as with c-erbB-2, there may be situations where parts of the receptor are "shed" into the bloodstream. In this case, the presence of the
20 receptor could be detected, for example, by use of an ELISA assay, and this would be a useful diagnostic tool for the identification of patients predisposed to develop or liable to develop tumours, in particular breast cancer.

25 Accordingly the present invention provides a method for detecting a tyrosine kinase having an amino acid sequence as defined in

SEQ ID NO. 1

or a fragment thereof which comprises reacting a test sample
30 with a specific antibody raised against an antigen from the said amino acid sequence and determining whether there is any antigen-antibody binding within the test sample. The test sample may be for example a tissue sample, such as a tumour sample, or a blood sample and presence of abnormal amounts of
35 the protein tyrosine kinase may indicate development of or susceptibility to develop a tumour.

The invention also provides a method for the detection in a sample of DNA a nucleic acid sequence encoding a tyrosine

kinase which comprises subjecting the sample or nucleic acid isolated therefrom to a method capable of detecting a nucleic acid sequence therein identical or substantially homologous to a defined nucleic acid sequence and identifying any
5 sequence identical or substantially homologous to all or part of the DNA sequence defined in

SEQ ID NO. 1,
or mRNA which would be the transcription product of such DNA. Again, presence of abnormal amounts of nucleic acid encoding
10 the protein tyrosine kinase may indicate development of or susceptibility to develop a tumour.

Suitable methods for the detection of specific DNA sequences include Southern blotting and/or PCR using
15 appropriate primers. Suitable methods for detecting mRNA include Northern blotting, RNase protection studies and direct nucleic acid *in situ* hybridisation.

EXAMPLE

20 Experimental work on which the invention is based is described in more detail in this Example in which reference is made to the accompanying drawings in which:

25 Figure 1 shows a Northern blot to examine expression of DDR mRNA in human axillary lymph nodes. 5µg of total RNA were loaded per lane. Approximately 4 kb DDR transcript (upper panel). Ethidium bromide staining of 28S rRNA, to demonstrate RNA loading, is shown (bottom panel).

30 Figure 2 shows *in situ* hybridisation to examine DDR expression in an involved lymph node. Light and dark field images are on the left and right respectively. Specific probe (DDR) and control (Cont.) are shown. Oligonucleotide probes were used.
35 Magnification x 200.

Figure 3 shows a Northern blot to examine DDR expression in human breast tissue. 10µg of total RNA per lane. Upper panel shows DDR transcript, lower panel ethidium bromide stained 28S

rRNA.

Figure 4 shows DDR expression in murine mammary development. 2µg Poly (A)+ RNA per lane. DDR transcript shown in upper panel. Blot was hybridised with GAPDH to demonstrate RNA loading (lower panel).

Figure 5 shows Northern blot analysis of DDR expression in human cell lines performed as for Figure 1. DDR transcript (upper panel), ethidium bromide stained 28S rRNA (lower panel).

Figure 6 shows *in situ* hybridisation analysis of DDR expression in human breast tumour tissue. Tissues were from three different individuals: a, b and c. In each panel, light and dark field images are on the left and right respectively. Specific probes are shown (DDR), and controls (Cont.). il= infiltrating lymphocytes, tu= tumour and ne= normal epithelium. Experiments 6a and 6b used oligonucleotides probes, and 6c used RNA probes. Magnification x 200.

Figure 7 shows DDR expression in normal tissues. a) human multiple tissue Northern blot containing 2µg poly (A)+ RNA per lane. DDR transcript shown in upper panel. Blot was stripped and rehybridised with human GAPDH probe to demonstrate RNA loading (lower panel). b) Total RNA Northern blot, 10µg RNA per lane, DDR transcript upper panel, ethidium bromide stained 28S rRNA lower panel.

MATERIALS AND METHODS

Tissues

Human tumour bearing or non-involved lymph nodes, tumour and normal tissue samples were obtained from surgery and rapidly frozen in liquid nitrogen. Normal breast tissue was obtained from reduction mammoplasty. Murine mammary tissue was obtained by excision of the fourth gland, removal of the associated lymph node, and rapid freezing in liquid nitrogen. All tissues were stored at -70°C or in liquid nitrogen.

Histopathological examination determined the presence or absence of metastases in the lymph nodes.

Cell lines

- 5 CAL 51, a human breast carcinoma cell line, was isolated and described by Gioanni et al (Br. J. Cancer, 62, 8-13 (1990)). The fibrosarcoma cell line HT1080, and the other breast cancer cell lines were purchased from the American Type Culture Collection. Cells were cultured according to conditions
10 recommended by the suppliers.

RNA extraction

- Total RNA was extracted from all tissues and cell lines by the method of Chomczynski and Sacchi (Anal. Biochem., 162, 156-159
15 (1987)). Poly (A)+ RNA was isolated from murine mammary gland total RNA samples using the Poly A+ Tract Kit (Promega) according to the manufacturer's instructions, except that mRNA was eluted with three washes (2 x 200µl, 1 x 100µl) of water.

20 RT-PCR and PTK differential screen

- 5µg of total RNA purified from involved or non-involved lymph nodes was digested with 10 units of RNase free DNase 1 (Pharmacia) using the manufacturer's recommended conditions in the presence of 40 units of RNase inhibitor (Boehringer
25 Mannheim). Samples were heated to 100°C for 1 minute (min) to inactivate the DNase, extracted with phenol/chloroform and precipitated with ethanol. For cDNA synthesis, 2.5µg of RNA was resuspended in water and denatured at 65°C for 3 min in the presence of 40 ng pd(N)₆ random hexamers (Pharmacia) and
30 cooled on ice. Reverse transcription was carried out using 1 mM dNTPs (Pharmacia), 40 units RNase inhibitor, 30 units of AMV reverse transcriptase SL (Life Sciences) and the buffer provided by the manufacturer at 41°C for 1 hour. Control reactions were performed as above but omitting the reverse
35 transcriptase enzyme (RT-control). To inactivate the inhibitory properties of reverse transcriptase on Taq polymerase, samples were heated to 100°C for 1 min.

Degenerate oligonucleotides were designed to amplify by PCT

an approximately 200 bp region encoding part of the catalytic domain of PTKs. The highly conserved amino acid sequence shown in

SEQ ID NO. 3

5 was used for designing the 5' oligonucleotide shown in

SEQ ID NO. 4

and the highly conserved amino acid sequence shown in

SEQ ID NO. 5

was used for designing the 3' oligonucleotide shown in

10 SEQ ID NO. 6.

Amplifications were carried out using cDNA derived from 0.5µg of original RNA in 60 mM KCl, 15mM Tris-HCl, 1.1mM MgCl₂, 200µM of each dNTP, 0.5ng of each primer and 2 units of Taq polymerase (Cetus). The reaction mixtures were cycled 30 times at 94°C for 1 min, 55°C for 2 min and 72°C for 1 min. For each sample, PCR was carried out on cDNA and the relevant RT-control. The amplified products were electrophoresed through 3% agarose gels, the approximately 200 bp band excised and purified using a Mermaid kit (Bio 101). PCR products from 20 involved nodes were subcloned into the pCR II vector (Invitrogen) according to the manufacturer's instructions. Clones were picked, grown in liquid culture, pelleted, resuspended in denaturing solution (0.5M NaOH, 1.5M NaCl) and aliquots were gridded onto duplicate Hybond-N membranes 25 (Amersham). Filters were then neutralised in 3M NaCl, 0.5M Tris-HCl (pH 7.4). Duplicate filters were hybridised with ³²P labelled randomly primed (Feinberg & Vogelstein, Anal. Biochem, 132, 6-13 (1983); and Feinberg & Vogelstein, Anal. Biochem, 137, 266-267 (1984)) PCR products from either the 30 involved or non-involved node. Clones containing PTK fragments that hybridised more strongly with the involved node-derived probe compared with the non-involved node derived probe were chosen for sequence analysis.

35 DNA sequencing

Templates for sequencing were prepared by the rapid method of Yie et al (Nucleic Acids Res., 21, 361(1993)) and were sequenced by the dideoxy method of Sanger et al, (Proc. Natl. Acad. Sci. (USA), 74, 5463-5467 (1977)) using a Sequenase

Version 2.0 kit (United States Biochemicals).

3' RACE-PCR

To isolate cDNA sequences 3' to the region amplified by the PTK PCR, a modification of the RACE protocol described by Frohman et al (Proc. Natl. Acad. Sci. (USA), 85, 8998-9002 (1988)) was used. RNA from human breast carcinoma cell line MDA MG 468 was reverse transcribed using the tagged oligo d(T) primer shown in

SEQ ID NO. 7.

Two nested sense oligonucleotides were used in consecutive PCR reactions. The upstream sense primer was as shown in

SEQ ID NO. 8

and the downstream sense primer was as shown in

SEQ ID NO. 9.

The tag sequence used with oligo d(T) for the reverse transcription was used as the antisense primer in both amplifications. The reaction buffer was as described for the PTK PCR except $MgCl_2$ was used at 1.5mM, cycling parameters were 94°C for 1 min 15 seconds, 56°C for 2 min and 72°C for 2 min.

Northern blots

Human total RNA samples were electrophoresed through formaldehyde/agarose gels, transferred to Hybond-N membranes (Amersham) and fixed by U.V. illumination. The human multiple tissue Northern blot was purchased from Clontech. Murine Poly (A)+ RNA samples were fractionated in agarose gels containing glyoxal/DMSO, transferred to ZetaProbe membranes (Biorad) and fixed by baking at 80°C for 2 hours. ^{32}P labelled probes were prepared as described by Feinberg & Vogelstein (supra). Hybridisations were carried out according to the manufacturer's instructions, membranes were washed to a final stringency of 0.1 x SSC, 0.1% SDS at 65°C.

35

Probe generation for *in situ* hybridisation

Riboprobes: The DDR RACE PCR fragment was sub-cloned into pBluescript KS⁺ (Stratagene) and sense and antisense RNAs labelled with ^{35}S -UTP (Amersham) were synthesised. The

transcription reactions were carried out using a Stratagene Transcription kit according to the manufacturer's instruction. Probes were subject to limited alkaline hydrolysis at 65°C for 100 min (Cox et al., Developmental Biology, 101, 485-502 (1984)) to achieve a final length of between 100-200 bases. Full length and hydrolysed probes were examined on formaldehyde/agarose gels.

Oligonucleotide probes: Antisense 30mer oligonucleotides as follows were synthesised to the 3' untranslated and catalytic domain encoding regions of DDR:-

- SEQ ID NO. 10
- SEQ ID NO. 11
- SEQ ID NO. 12
- SEQ ID NO. 13.

A 30mer random oligonucleotide N₃₀ was also synthesised (randomer).

The oligonucleotides were 3'-labelled with ³³P-dATP (Amersham) using terminal transferase according to manufacturer's instructions (Boehringer Mannheim). Probes were purified by passing through Chroma-spin 10 columns (Clontech) to remove unincorporated label. Tail lengths were analysed on a DNA sequencing gel. The four specific oligonucleotides were then mixed equally to form a cocktail.

Labelling efficiency of all probes was assessed by scintillation counting.

In situ hybridisation

Except those used in post-hybridisation steps, all solutions were diethylpyrocarbonate treated and all glassware was baked.

7µm cryostat sections were cut from snap frozen human tumour material which had been excised during surgery, and stored in either liquid nitrogen or at -80°C. Sections were dried onto APTES (3-aminopropyltriethoxysilane) treated slides (Rentrop et al., Histochem. J., 18, 271-276 (1986)) over dry ice, and fixed at 4°C in 4% paraformaldehyde, phosphate buffered saline

(PBS) for 15 min, followed by two 15 min washes in PBS. The sections were dehydrated through ethanols from 30%-100%, air dried, and stored at -80°C until required.

5 a) With riboprobes

Sections were equilibrated to room temperature for 5 min, quickly rehydrated through ethanols from 100% to 30%, and washed twice for 5 min in PBS. Sections were acetylated for 10 min in 0.1M triethanolamine, 0.25% acetic anhydride while stirring. After washing in 2xSSC for 2 min, denaturation of tissue RNAs was carried out in 50% formamide, 2xSSC for 10 min at 60°C. Sections were then dehydrated and air dried.

15 Hybridisation was performed overnight in buffer (50% formamide, 0.3M NaCl, 10mM Tris- HCl pH8, 2mM EDTA pH8, 10% dextran sulphate, 10mM sodium phosphate pH8, 50µgml⁻¹ssDNA) with comparable counts per minute (cpm) of either antisense or sense riboprobes, in a humid chamber at 50°C.

20 After hybridisation, two 30 min washes in 2xSSC, 50% formamide at 60°C were followed by treatment with 20µgml⁻¹ RNase A in 0.5M NaCl, 10mM Tris HCl pH7.5, 5mM EDTA pH 7.5 for 30 min at 37°C. Sections were washed for 15 min in 2xSSC at 60°C, followed by two final washes at a stringency of 0.1xSSC for 30 min at 60°C. Sections were dehydrated and air dried before autoradiography.

b) With oligonucleotide probes.

30 Sections were equilibrated to room temperature, rehydrated, washed and acetylated as described for sections hybridised with riboprobes. After washing twice for 5 min with PBS they were dehydrated and air dried.

35 Sections were overlaid with 10µl of hybridisation buffer (6x SSC, 50% formamide, 5x Denhardt's solution (Sambrook et al., Molecular Cloning: A Laboratory Manual. 2nd Edn., Cold Spring Harbor Laboratory Press, USA (1989)), 10% dextran sulphate)

containing $20\mu\text{gml}^{-1}$ unlabelled random oligonucleotide, and prehybridised for 3 hours at 37°C in a sealed humid chamber. After prehybridisation, a further $10\mu\text{l}$ of hybridisation buffer containing comparable cpm of either the labelled
5 oligonucleotide cocktail or labelled randomer were added to each section and gently mixed with prehybridisation solution. The sections were incubated overnight at 37°C (Lathe, *In Situ Hybridisation Principles and Practice*, Polak & McGee (eds.) Oxford Scientific Publications, Oxford pp71-80 (1989)) in a
10 sealed humid chamber.

Sections were removed from the chamber, and the hybridisation solution was rinsed off with prewarmed $5\times\text{SSC}$ gently pipetted over the slide. The area around the sections was dried and
15 the sections were then overlaid with $20\mu\text{l}$ prehybridisation solution ($20\mu\text{gml}^{-1}$ randomer) and incubated for 60 min as before. This process might help reduce background by replacing non-specifically bound oligonucleotide with random oligonucleotide. This step was repeated once. Slides were
20 washed to a final stringency of 60% formamide, $5\times\text{SSC}$ at 37°C . Sections were dehydrated in 70% and 100% ethanol and air dried.

Autoradiography on tissues

25 This process was the same for both riboprobe and oligonucleotide methods. In order to give a guideline of the length of time of exposure to photographic emulsion required, slides were placed against autoradiography film overnight. Slides were dipped in K.5 (Ilford) photographic emulsion
30 diluted 1:1 with 2% glycerol, following manufacturer's guidelines. Slides were dried in an air-tight container with desiccant at room temperature for 2 hours and stored dry at 4°C , in a light-tight box for the required length of time. Slides were developed in D19 (Kodak), stopped in 1% acetic
35 acid, 1% glycerol and fixed in 30% sodium thiosulphate. Sections were counter stained in 0.02% toluidine blue, mounted in DPX (Fisons) and examined by light and darkfield microscopy.

RESULTS

Identification of PTKs in a differential screen of involved and non-involved lymph nodes

5 RNA was isolated from involved and non-involved axillary lymph nodes from patients with metastatic breast tumours. RT-PCR was performed on the RNA using degenerate oligonucleotides encoding the amino acid sequences shown in SEQ ID NO. 3 and
10 SEQ ID NO 5, sequences highly conserved in PTKs (Hanks, Science, 241, 42-52, (1988)). The resultant pools of approximately 200bp PTK cDNA fragments from the involved nodes were subcloned, and clones screened with mixed pools of radiolabelled PTK fragments from both involved and non-involved nodes. Clones amplified preferentially from the
15 involved nodes were sequenced. For each of five involved lymph nodes, 50 or 100 clones were screened; differentially amplified cDNAs were found in four of five nodes. A novel kinase was detected as the major differentially amplified PTK-encoding cDNA; it showed similarities with trk A (Martin-Zanca
20 et al, Nature, 319, 743-748 (1986)) and the rat PCR generated fragment tyro 10 (Lai & Lemke, Neuron, 6, 691-704 (1991)). The receptor PTK c-erb-B2 was found at high levels in three nodes.

25 The trk A/tyro 10-like PTK is identical to DDR
To study the expression patterns of the trk-like kinase, a 3' Rapid Amplification of cDNA Ends (RACE)-PCR fragment was generated using two specific nested oligonucleotides
30 corresponding to the sequence between the two degenerate primers. Dideoxy sequencing of this 1.2 kilo base (kb) 3' RACE-PCR fragment revealed 384 nucleotides of predicted coding sequence. This comprised sequences encoding part of a putative PTK catalytic domain, which showed similarity to trk A, followed by a short C terminal domain of 8 amino acids
35 after subdomain XI. While analysing this fragment, a report was published describing a receptor PTK found in cultured breast carcinoma cells lines with an extracellular discoidin 1-like domain (DDR) (Johnson et al, Proc. Natl. Acad. Sci. (USA), 90, 5677-5681 (1993)). Comparison of the DDR sequence

with the 3' RACE-PCR fragment revealed them to be identical, the 3' RACE-PCR fragment begins at nucleotide position 2500 of the published DDR sequence.

5 Expression of DDR mRNA in lymph nodes

DDR was detected in a PCR based differential screen designed to identity PTKs expressed at higher levels in tumour containing compared with tumour free lymph nodes. In order to examine whether higher levels of DDR mRNA were present in the involved compared with the non-involved lymph node, the 3' RACE-PCR fragment radiolabelled with ^{32}P was used as a probe to hybridise with Northern blots of lymph node RNA. This probe recognised a single band of approximately 4kb on Northern blots. DDR mRNA was undetectable in the non-involved lymph node and present in four of six metastasis-containing nodes (Figure 1). In situ hybridisation on a section of involved lymph node showed specific hybridisation over the metastasised tumour cells (Figure 2).

20 Expression of DDR mRNA in breast tumours, breast carcinoma cell lines and mammary development

The expression patterns of DDR in a panel of breast tumour RNAs and a number of breast carcinoma cell lines were examined by hybridising Northern blots with the ^{32}P labelled 3' RACE-PCR fragment. DDR mRNA was present in normal breast tissue and in all the breast tumours tested (figure 3). Elevated levels of expression were detected in samples of 'normal' breast and carcinoma, both surgically removed from a pregnant woman. To examine DDR expression in mammary gland development, a Northern blot of murine mammary gland poly (A)+ RNA was hybridised with the 3' RACE-PCR derived probe. Figure 4 shows that murine DDR mRNA (NEP, Zerlin *et al*, Oncogene, 8, 2731-2739 (1983)), see discussion) was detected in all stages of development. Expression was moderate in virgin glands, a slight increase was seen in mid-pregnancy, during lactation very low levels of DDR mRNA were detected, and moderate levels reappeared during involution. During lactation casein mRNA comprises a large proportion of the total, this effectively dilutes the levels of other mRNAs; glyceraldehyde 3-phosphate

dehydrogenase (GAPDH) mRNA levels were decreased in these samples also. DDR mRNA expression was also detected in fat pad samples. The fat pad represents mammary tissue which has developed in the absence of epithelial cells due to the cauterisation of the nipple and surrounding tissue prior to epithelial outgrowth. Varying amounts of DDR mRNA were detected in the breast cancer cell lines examined, the highest were in MCF 7 and T47D and the lowest were in MDA MB 157 and BT 474. Expression was not detected in the fibrosarcoma cell line HT1080 (Figure 5).

Northern blot analysis revealed that DDR mRNA was present in normal breast and breast tumour tissue. *In situ* hybridisation was, therefore, performed on normal breast tissue and primary breast carcinoma samples to determine which cell types express DDR. In sections from three tumour samples, (Figures 6 a, b and c) hybridisation was predominantly over the tumour cells. DDR mRNA was undetectable in normal breast tissue adjacent to the tumour (Figure 6c top panel), and in all normal breast samples examined (data not shown). Figure 6c (top panel), shows a mass of invasive carcinoma surrounded by stromal cells in the top right of the field, with normal epithelial tissue in the bottom left of the field; hybridisation was specifically over the tumour cells, and DDR mRNA was undetectable in the normal epithelium and stromal cells. Figure 6c (middle panel) shows DDR positive cancer cells which have replaced the normal epithelial cells in the breast lobules. The malignant epithelial elements are surrounded by a dense lymphocytic infiltrate in which DDR was undetectable. A comparable result to Figure 6c was obtained using oligonucleotide probes, (date not shown). Figures 6a and c (lower panels) show results obtained with the random oligonucleotide probe and the sense riboprobe respectively, with no specific hybridisation. These controls were carried out for all *in situ* hybridisations (not shown in all cases).

Patterns of DDR expression in normal human tissues

A human poly (A)+ RNA multiple tissue Northern blot was screened with the 3' RACE-PCR probe (Figure 7a). DDR was

expressed at high levels in kidney and placental tissue, low levels were detected in the heart and lung, and expression was undetectable in the liver. A more extensive panel of human tissues was examined on total RNA Northern blots (Figure 7b).

5 DDR was expressed in a wide range of tissues, with relatively high levels in the kidney cortex and thyroid. By this analysis, the haematopoietic tissues bone marrow and spleen did not have detectable levels of DDR mRNA.

10 DISCUSSION

The initial differential screen identified DDR as a PTK expressed in involved but not in non-involved lymph nodes from breast cancer patients. Further studies revealed varying

15 levels of DDR mRNA in both normal breast and all of the breast carcinomas analysed. Elevated mRNA levels, relative to normal breast tissue, were found in both apparently normal breast tissue and breast carcinoma taken from the same pregnant patient. However, an examination of murine mammary gland

20 development revealed only a moderate increase in DDR mRNA in pregnancy. The higher levels of DDR in the human pregnant breast could be due to a focus of infiltrating carcinoma, as the specimen was taken from the tissue adjacent to a tumour. Alternatively, the murine mammary gland may differ from human

25 with respect to DDR expression during pregnancy. Unfortunately material was not available from the human pregnant breast for histopathological examination or *in situ* hybridisation studies to resolve this issue. *In situ* hybridisation studies were performed to determine which cells

30 in a particular tissue sample express DDR mRNA, and to give an indication of the relative abundance of mRNA expressed by different cell types. *In situ* hybridisation of a heavily infiltrated lymph node revealed DDR mRNA at high levels in the metastasising tumour cells and undetectable levels in the

35 adjacent cells. DDR mRNA was also shown to be present in tumour cells in primary breast carcinomas by *in situ* hybridisation. Where a direct comparison could be made, high levels of DDR mRNA were detected in the tumour cells but not in adjacent normal epithelial cells nor in tumour infiltrating

lymphocytes. By Northern analysis, DDR mRNA was present in normal human breast tissue, but by *in situ* hybridisation on human mammary tissue DDR has only been detected in tumour cells, which are epithelial in origin, and not in normal human breast epithelial or stromal cells. The question of the site of DDR expression in normal human mammary tissue is therefore unanswered; most likely cellular mRNA levels are too low for detection by *in situ* hybridisation. Nonetheless over expression of DDR occurs in some breast tumour cells when compared with normal mammary epithelium. Having demonstrated the presence of DDR mRNA in several sources of breast tumour tissue (in primary tumours, metastatic cells and carcinoma-derived cell lines) the expression in normal tissues was examined by Northern blotting. A wide range of levels were observed, ranging from high in the kidney cortex and thyroid, moderate in breast and lung, to undetectable amounts in the bone marrow, spleen, and liver. DDR expression is not restricted to any one tissue type, therefore, its role is not tissue specific. However, it is not ubiquitously expressed, thus a housekeeping function is unlikely and a cell specific role is probable.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:

(A) NAME: The Institute of Cancer Research
 (B) STREET: 17A, Onslow Gardens
 (C) CITY: London
 (E) COUNTRY: Great Britain
 (F) POSTAL CODE (ZIP): SW7 3AL

(A) NAME: The Wellcome Foundation Limited
 (B) STREET: Unicorn House, 160 Euston Road
 (C) CITY: London
 (E) COUNTRY: Great Britain
 (F) POSTAL CODE (ZIP): NW1 2BP

(ii) TITLE OF INVENTION: Cell Growth Factor Receptors

(iii) NUMBER OF SEQUENCES: 13

(iv) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk
 (B) COMPUTER: IBM PC compatible
 (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 (D) SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 3754 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

(A) NAME/KEY: CDS
 (B) LOCATION: 142..2883

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

```

CCCCGGGTCGG ACCGCCTGGG TCTGCCGGGA AGAGCGATGA GAGGTGTCTG AAGGTGGCTA      60
TTCACGTGACC GATGGGGTTG GACTTGAAGG AATGCCAAGA GATGCTGCCC CCACCCCCTT      120
AGGCCCCGAGG GATCAGGAGC T ATG GGA CCA GAG GCC CTG TCA TCT TTA CTG      171
                               Met Gly Pro Glu Ala Leu Ser Ser Leu Leu
                               1                               5                               10
CTG CTG CTC TTG GTG GCA AGT GGA GAT GCT GAC ATG AAG GGA CAT TTT      219
Leu Leu Leu Leu Val Ala Ser Gly Asp Ala Asp Met Lys Gly His Phe
                               15                               20                               25
GAT CCT GCC AAG TGC CGC TAT GCC CTG GGC ATG CAG GAC CGG ACC ATC      267
Asp Pro Ala Lys Cys Arg Tyr Ala Leu Gly Met Gln Asp Arg Thr Ile
                               30                               35                               40
CCA GAC AGT GAC ATC TCT GCT TCC AGC TCC TGG TCA GAT TCC ACT GCC      315
Pro Asp Ser Asp Ile Ser Ala Ser Ser Ser Trp Ser Asp Ser Thr Ala
                               45                               50                               55

```

	GCC CGC CAC AGC AGG TTG GAG AGC AGT GAC GGG GAT GGG GCC TGG TGC	363
	Ala Arg His Ser Arg Leu Glu Ser Ser Asp Gly Asp Gly Ala Trp Cys	
	60 65 70	
5	CCC GCA GGG TCG GTG TTT CCC AAG GAG GAG GAG TAC TTG CAG GTG GAT	411
	Pro Ala Gly Ser Val Phe Pro Lys Glu Glu Glu Tyr Leu Gln Val Asp	
	75 80 85 90	
10	CTA CAA CGA CTC CAC CTG GTG GCT CTG GTG GGC ACC CAG GGA CGG CAT	459
	Leu Gln Arg Leu His Leu Val Ala Leu Val Gly Thr Gln Gly Arg His	
	95 100 105	
15	GCC GGG GGC CTG GGC AAG GAG TTC TCC CGG AGC TAC CGG CTG CGT TAC	507
	Ala Gly Gly Leu Gly Lys Glu Phe Ser Arg Ser Tyr Arg Leu Arg Tyr	
	110 115 120	
20	TCC CGG GAT GGT CGC CGC TGG ATG GGC TGG AAG GAC CGC TGG GGT CAG	555
	Ser Arg Asp Gly Arg Arg Trp Met Gly Trp Lys Asp Arg Trp Gly Gln	
	125 130 135	
25	GAG GTG ATC TCA GGC AAT GAG GAC CCT GAG GGA GTG GTG CTG AAG GAC	603
	Glu Val Ile Ser Gly Asn Glu Asp Pro Glu Gly Val Val Leu Lys Asp	
	140 145 150	
30	CTT GGG CCC CCC ATG GTT GCC CGA CTG GTT CGC TTC TAC CCC CGG GCT	651
	Leu Gly Pro Pro Met Val Ala Arg Leu Val Arg Phe Tyr Pro Arg Ala	
	155 160 165 170	
35	GAC CGG GTC ATG AGC GTC TGT CTG CGG GTA GAG CTC TAT GGC TGC CTC	699
	Asp Arg Val Met Ser Val Cys Leu Arg Val Glu Leu Tyr Gly Cys Leu	
	175 180 185	
40	TGG AGG GAT GGA CTC CTG TCT TAC ACC GCC CCT GTG GGG CAG ACA ATG	747
	Trp Arg Asp Gly Leu Leu Ser Tyr Thr Ala Pro Val Gly Gln Thr Met	
	190 195 200	
45	TAT TTA TCT GAG GCC GTG TAC CTC AAC GAC TCC ACC TAT GAC GGA CAT	795
	Tyr Leu Ser Glu Ala Val Tyr Leu Asn Asp Ser Thr Tyr Asp Gly His	
	205 210 215	
50	ACC GTG GGC GGA CTG CAG TAT GGG GGT CTG GGC CAG CTG GCA GAT GGT	843
	Thr Val Gly Gly Leu Gln Tyr Gly Gly Leu Gly Gln Leu Ala Asp Gly	
	220 225 230	
55	GTG GTG GGG CTG GAT GAC TTT AGG AAG AGT CAG GAG CTG CGG GTC TGG	891
	Val Val Gly Leu Asp Asp Phe Arg Lys Ser Gln Glu Leu Arg Val Trp	
	235 240 245 250	
60	CCA GGC TAT GAC TAT GTG GGA TGG AGC AAC CAC AGC TTC TCC AGT GGC	939
	Pro Gly Tyr Asp Tyr Val Gly Trp Ser Asn His Ser Phe Ser Ser Gly	
	255 260 265	
65	TAT GTG GAG ATG GAG TTT GAG TTT GAC CGG CTG AGG GCC TTC CAG GCT	987
	Tyr Val Glu Met Glu Phe Glu Phe Asp Arg Leu Arg Ala Phe Gln Ala	
	270 275 280	
70	ATG CAG GTC CAC TGT AAC AAC ATG CAC ACG CTG GGA GCC CGT CTG CCT	1035
	Met Gln Val His Cys Asn Asn Met His Thr Leu Gly Ala Arg Leu Pro	
	285 290 295	
75	GGC GGG GTG GAA TGT CGC TTC CGG CGT GGC CCT GCC ATG GCC TGG GAG	1083
	Gly Gly Val Glu Cys Arg Phe Arg Arg Gly Pro Ala Met Ala Trp Glu	
	300 305 310	
80	GGG GAG CCC ATG CGC CAC AAC CTA GGG GGC AAC CTG GGG GAC CCC AGA	1131
	Gly Glu Pro Met Arg His Asn Leu Gly Gly Asn Leu Gly Asp Pro Arg	
	315 320 325 330	

	GCC	CGG	GCT	GTC	TCA	GTG	CCC	CTT	GGC	GGC	CGT	GTG	GCT	CGC	TTT	CTG	1179
	Ala	Arg	Ala	Val	Ser	Val	Pro	Leu	Gly	Gly	Arg	Val	Ala	Arg	Phe	Leu	
					335					340					345		
5	CAG	TGC	CGC	TTC	CTC	TTT	GCG	GGG	CCC	TGG	TTA	CTC	TTC	AGC	GAA	ATC	1227
	Gln	Cys	Arg	Phe	Leu	Phe	Ala	Gly	Pro	Trp	Leu	Leu	Phe	Ser	Glu	Ile	
				350					355					360			
10	TCC	TTC	ATC	TCT	GAT	GTG	GTG	AAC	AAT	TCC	TCT	CCG	GCA	CTG	GGA	GGC	1275
	Ser	Phe	Ile	Ser	Asp	Val	Val	Asn	Asn	Ser	Ser	Pro	Ala	Leu	Gly	Gly	
			365					370						375			
15	ACC	TTC	CCG	CCA	GCC	CCC	TGG	TGG	CCG	CCT	GGC	CCA	CCT	CCC	ACC	AAC	1323
	Thr	Phe	Pro	Pro	Ala	Pro	Trp	Trp	Pro	Pro	Gly	Pro	Pro	Pro	Thr	Asn	
		380					385					390					
20	TTC	AGC	AGC	TTG	GAG	CTG	GAG	CCC	AGA	GGC	CAG	CCA	AGG	CCC	GTG	GCC	1371
	Phe	Ser	Ser	Leu	Glu	Leu	Glu	Pro	Arg	Gly	Gln	Pro	Arg	Pro	Val	Ala	
	395				400					405					410		
25	AAG	GCC	GAG	GGG	AGC	CCG	ACC	GCC	ATC	CTC	ATC	GGC	TGC	CTG	GTG	GCC	1419
	Lys	Ala	Glu	Gly	Ser	Pro	Thr	Ala	Ile	Leu	Ile	Gly	Cys	Leu	Val	Ala	
					415					420				425			
30	ATC	ATC	CTG	CTC	CTG	CTG	CTC	ATC	ATT	GCC	CTC	ATG	CTC	TGG	CGG	CTG	1467
	Ile	Ile	Leu	Leu	Leu	Leu	Leu	Ile	Ile	Ala	Leu	Met	Leu	Trp	Arg	Leu	
				430				435						440			
35	CAC	TGG	CGC	AGG	CTC	CTC	AGC	AAG	GCT	GAA	CGG	AGG	GTG	TTG	GAA	GAG	1515
	His	Trp	Arg	Arg	Leu	Leu	Ser	Lys	Ala	Glu	Arg	Arg	Val	Leu	Glu	Glu	
			445					450					455				
40	GAG	CTG	ACG	GTT	CAC	CTC	TCT	GTC	CCT	GGG	GAC	ACT	ATC	CTC	ATC	AAC	1563
	Glu	Leu	Thr	Val	His	Leu	Ser	Val	Pro	Gly	Asp	Thr	Ile	Leu	Ile	Asn	
	460						465					470					
45	AAC	CGC	CCA	GGT	CCT	AGA	GAG	CCA	CCC	CCG	TAC	CAG	GAG	CCC	CGG	CCT	1611
	Asn	Arg	Pro	Gly	Pro	Arg	Glu	Pro	Pro	Pro	Tyr	Gln	Glu	Pro	Arg	Pro	
	475					480					485				490		
50	CGT	GGG	AAT	CCG	CCC	CAC	TCC	GCT	CCC	TGT	GTC	CCC	AAT	GGC	TCT	GCG	1659
	Arg	Gly	Asn	Pro	Pro	His	Ser	Ala	Pro	Cys	Val	Pro	Asn	Gly	Ser	Ala	
				495						500				505			
55	TTG	CTG	CTC	TCC	AAT	CCA	GCC	TAC	CGC	CTC	CTT	CTG	GCC	ACT	TAC	GCC	1707
	Leu	Leu	Leu	Ser	Asn	Pro	Ala	Tyr	Arg	Leu	Leu	Leu	Ala	Thr	Tyr	Ala	
				510				515						520			
60	CGT	CCC	CCT	CGA	GGC	CCG	GGC	CCC	CCC	ACA	CCC	GCC	TGG	GCC	AAA	CCC	1755
	Arg	Pro	Pro	Arg	Gly	Pro	Gly	Pro	Pro	Thr	Pro	Ala	Trp	Ala	Lys	Pro	
			525					530					535				
65	ACC	AAC	ACC	CAG	GCC	TAC	AGT	GGG	GAC	TAT	ATG	GAG	CCT	GAG	AAG	CCA	1803
	Thr	Asn	Thr	Gln	Ala	Tyr	Ser	Gly	Asp	Tyr	Met	Glu	Pro	Glu	Lys	Pro	
		540					545					550					
65	GGC	GCC	CCG	CTT	CTG	CCC	CCA	CCT	CCC	CAG	AAC	AGC	GTC	CCC	CAT	TAT	1851
	Gly	Ala	Pro	Leu	Leu	Pro	Pro	Pro	Pro	Gln	Asn	Ser	Val	Pro	His	Tyr	
	555				560					565					570		
65	GCC	GAG	GCT	GAC	ATT	GTT	ACC	CTG	CAG	GGC	GTC	ACC	GGG	GGC	AAC	ACC	1899
	Ala	Glu	Ala	Asp	Ile	Val	Thr	Leu	Gln	Gly	Val	Thr	Gly	Gly	Asn	Thr	
				575						580				585			
65	TAT	GCT	GTG	CCT	GCA	CTG	CCC	CCA	GGG	GCA	GTC	GGG	GAT	GGG	CCC	CCC	1947
	Tyr	Ala	Val	Pro	Ala	Leu	Pro	Pro	Gly	Ala	Val	Gly	Asp	Gly	Pro	Pro	
				590					595					600			

	AGA	GTG	GAT	TTC	CCT	CGA	TCT	CGA	CTC	CGC	TTC	AAG	GAG	AAG	CTT	GGC	1995
	Arg	Val	Asp	Phe	Pro	Arg	Ser	Arg	Leu	Arg	Phe	Lys	Glu	Lys	Leu	Gly	
			605					610					615				
5	GAG	GGC	CAG	TTT	GGG	GAG	GTG	CAC	CTG	TGT	GAG	GTC	GAC	AGC	CCT	CAA	2043
	Glu	Gly	Gln	Phe	Gly	Glu	Val	His	Leu	Cys	Glu	Val	Asp	Ser	Pro	Gln	
		620					625					630					
10	GAT	CTG	GTT	AGT	CTT	GAT	TTC	CCC	CTT	AAT	GTG	CGT	AAG	GGA	CAC	CCT	2091
	Asp	Leu	Val	Ser	Leu	Asp	Phe	Pro	Leu	Asn	Val	Arg	Lys	Gly	His	Pro	
	635					640					645					650	
15	TTG	CTG	GTA	GCT	GTC	AAG	ATC	TTA	CGG	CCA	GAT	GCC	ACC	AAG	AAT	GCC	2139
	Leu	Leu	Val	Ala	Val	Lys	Ile	Leu	Arg	Pro	Asp	Ala	Thr	Lys	Asn	Ala	
					655					660					665		
20	AGG	AAT	GAT	TTC	CTG	AAA	GAG	GTG	AAG	ATC	ATG	TCG	AGG	CTC	AAG	GAC	2187
	Arg	Asn	Asp	Phe	Leu	Lys	Glu	Val	Lys	Ile	Met	Ser	Arg	Leu	Lys	Asp	
				670					675					680			
25	CCA	AAC	ATC	ATT	CGG	CTG	CTG	GGC	GTG	TGT	GTG	CAG	GAC	GAC	CCC	CTC	2235
	Pro	Asn	Ile	Ile	Arg	Leu	Leu	Gly	Val	Cys	Val	Gln	Asp	Asp	Pro	Leu	
			685					690					695				
30	TGC	ATG	ATT	ACT	GAC	TAC	ATG	GAG	AAC	GGC	GAC	CTC	AAC	CAG	TTC	CTC	2283
	Cys	Met	Ile	Thr	Asp	Tyr	Met	Glu	Asn	Gly	Asp	Leu	Asn	Gln	Phe	Leu	
		700					705					710					
35	AGT	GCC	CAC	CAG	CTG	GAG	GAC	AAG	GCA	GCC	GAG	GGG	GCC	CCT	GGG	GAC	2331
	Ser	Ala	His	Gln	Leu	Glu	Asp	Lys	Ala	Ala	Glu	Gly	Ala	Pro	Gly	Asp	
		715				720					725				730		
40	GGG	CAG	GCT	GCG	CAG	GGG	CCC	ACC	ATC	AGC	TAC	CCA	ATG	CTG	CTG	CAT	2379
	Gly	Gln	Ala	Ala	Gln	Gly	Pro	Thr	Ile	Ser	Tyr	Pro	Met	Leu	Leu	His	
					735					740				745			
45	GTG	GCA	GCC	CAG	ATC	GCC	TCC	GGC	ATG	CGC	TAT	CTG	GCC	ACA	CTC	AAC	2427
	Val	Ala	Ala	Gln	Ile	Ala	Ser	Gly	Met	Arg	Tyr	Leu	Ala	Thr	Leu	Asn	
				750					755					760			
50	TTT	GTA	CAT	CGG	GAC	CTG	GCC	ACG	CGG	AAC	TGC	CTA	GTT	GGG	GAA	AAT	2475
	Phe	Val	His	Arg	Asp	Leu	Ala	Thr	Arg	Asn	Cys	Leu	Val	Gly	Glu	Asn	
			765					770					775				
55	TTC	ACC	ATC	AAA	ATC	GCA	GAC	TTT	GGC	ATG	AGC	CGG	AAC	CTC	TAT	GCT	2523
	Phe	Thr	Ile	Lys	Ile	Ala	Asp	Phe	Gly	Met	Ser	Arg	Asn	Leu	Tyr	Ala	
		780					785					790					
60	GGG	GAC	TAT	TAC	CGT	GTG	CAG	GGC	CGG	GCA	GTG	CTG	CCC	ATC	CGC	TGG	2571
	Gly	Asp	Tyr	Tyr	Arg	Val	Gln	Gly	Arg	Ala	Val	Leu	Pro	Ile	Arg	Trp	
		795				800					805				810		
65	ATG	GCC	TGG	GAG	TGC	ATC	CTC	ATG	GGG	AAG	TTC	ACG	ACT	GCG	AGT	GAC	2619
	Met	Ala	Trp	Glu	Cys	Ile	Leu	Met	Gly	Lys	Phe	Thr	Thr	Ala	Ser	Asp	
				815					820					825			
70	GTG	TGG	GCC	TTT	GGT	GTG	ACC	CTG	TGG	GAG	GTG	CTG	ATG	CTC	TGT	AGG	2667
	Val	Trp	Ala	Phe	Gly	Val	Thr	Leu	Trp	Glu	Val	Leu	Met	Leu	Cys	Arg	
				830					835					840			
75	GCC	CAG	CCC	TTT	GGG	CAG	CTC	ACC	GAC	GAG	CAG	GTC	ATC	GAG	AAC	GCG	2715
	Ala	Gln	Pro	Phe	Gly	Gln	Leu	Thr	Asp	Glu	Gln	Val	Ile	Glu	Asn	Ala	
			845					850					855				
80	GGG	GAG	TTC	TTC	CGG	GAC	CAG	GGC	CGG	CAG	GTG	TAC	CTG	TCC	CGG	CCG	2763
	Gly	Glu	Phe	Phe	Arg	Asp	Gln	Gly	Arg	Gln	Val	Tyr	Leu	Ser	Arg	Pro	
		860				865						870					

	CCT GCC TGC CCG CAG GGC CTA TAT GAG CTG ATG CTT CGG TGC TGG AGC	2811
	Pro Ala Cys Pro Gln Gly Leu Tyr Glu Leu Met Leu Arg Cys Trp Ser	
	875 880 885 890	
5	CGG GAG TCT GAG CAG CGA CCA CCC TTT TCC CAG CTG CAT CGG TTC CTG	2859
	Arg Glu Ser Glu Gln Arg Pro Pro Phe Ser Gln Leu His Arg Phe Leu	
	895 900 905	
10	GCA GAG GAT GCA CTC AAC ACG GTG TGAATCACAC ATCCAGCTGC CCCTCCCTCA	2913
	Ala Glu Asp Ala Leu Asn Thr Val	
	910	
	GGGAGCGATC CAGGGGAAGC CAGTGACACT AAAACAAGAG GACACAATGG CACCTCTGCC	2973
15	CCTTCCCCTC CCGACAGCCC ATCACCTCTA ATAGAGGCAG TGAGACTGCA GGCTGGGCCC	3033
	ACCCAGGGAG CTGATGCCCC TTCTCCCCTT CCTGGACACA CTCTCATGTC CCCTTCCTGT	3093
20	TCTTCCTTCC TAGAAGCCCC TGTCGCCCCAC CCAGCTGGTC CTGTGGATGG GATCCTCTCC	3153
	ACCCACCTCT AGCCATCCCT TGGGGAAGGG TGGGGAGAAA TATAGGATAG ACACTGGACA	3213
	TGGCCCATTG GAGCACCTGG GCCCCACTGG ACAACACTGA TTCCTGGACA GGTGGCTGCG	3273
25	CCCCAGCTT CTCTCTCCCT GTCACACACT GGACCCCACT GGCTGAGAAT CTGGGGGTGA	3333
	GGAGGACAAG AAGGAGAGGA AAATGTTTCC TTGTGCCTGC TCCTGTACTT GTCCTCAGCT	3393
30	TGGGCTTCTT CCTCCTCCAT CACCTGAAAC ACTGGACCTG GGGGTAGCCC CGCCCCAGCC	3453
	CTCAGTCACC CCCCCTTCC CACCTGCAGT CTTGTAGCTA GAACTTCTCT AAGCCTATAC	3513
	GTTTCTGTGG AGTAAATATT GGGATTGGGG GGAAAGAGGG AGCAACGGCC CATAGCCTTG	3573
35	GGGTTGGACA TCTCTAGTGT AGCTGCCACA TTGATTTTTTC TATAATCACT TGGGTTTGTA	3633
	CATTTTTTGGG GGGAGAGACA CAGATTTTTA CACTAATATA TGGACCTAGC TTGAGGCAAT	3693
40	TTTAATCCCC TGCACTAGGC AGGTAATAAT AAAGGTTGAG TTTTCCACAA AAAAAAAAAA	3753
	A	3754

(2) INFORMATION FOR SEQ ID NO: 2:

45

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 914 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

50

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

55	Met Gly Pro Glu Ala Leu Ser Ser Leu Leu Leu Leu Leu Val Ala	
	1 5 10 15	
	Ser Gly Asp Ala Asp Met Lys Gly His Phe Asp Pro Ala Lys Cys Arg	
	20 25 30	
60	Tyr Ala Leu Gly Met Gln Asp Arg Thr Ile Pro Asp Ser Asp Ile Ser	
	35 40 45	
65	Ala Ser Ser Ser Trp Ser Asp Ser Thr Ala Ala Arg His Ser Arg Leu	
	50 55 60	
	Glu Ser Ser Asp Gly Asp Gly Ala Trp Cys Pro Ala Gly Ser Val Phe	
	65 70 75 80	

31

	Pro	Lys	Glu	Glu	Glu	Tyr	Leu	Gln	Val	Asp	Leu	Gln	Arg	Leu	His	Leu	
					85					90					95		
5	Val	Ala	Leu	Val	Gly	Thr	Gln	Gly	Arg	His	Ala	Gly	Gly	Leu	Gly	Lys	
				100					105					110			
	Glu	Phe	Ser	Arg	Ser	Tyr	Arg	Leu	Arg	Tyr	Ser	Arg	Asp	Gly	Arg	Arg	
			115					120					125				
10	Trp	Met	Gly	Trp	Lys	Asp	Arg	Trp	Gly	Gln	Glu	Val	Ile	Ser	Gly	Asn	
		130					135					140					
	Glu	Asp	Pro	Glu	Gly	Val	Val	Leu	Lys	Asp	Leu	Gly	Pro	Pro	Met	Val	
15		145				150					155					160	
	Ala	Arg	Leu	Val	Arg	Phe	Tyr	Pro	Arg	Ala	Asp	Arg	Val	Met	Ser	Val	
					165					170					175		
20	Cys	Leu	Arg	Val	Glu	Leu	Tyr	Gly	Cys	Leu	Trp	Arg	Asp	Gly	Leu	Leu	
				180					185					190			
	Ser	Tyr	Thr	Ala	Pro	Val	Gly	Gln	Thr	Met	Tyr	Leu	Ser	Glu	Ala	Val	
			195					200					205				
25	Tyr	Leu	Asn	Asp	Ser	Thr	Tyr	Asp	Gly	His	Thr	Val	Gly	Gly	Leu	Gln	
		210					215					220					
	Tyr	Gly	Gly	Leu	Gly	Gln	Leu	Ala	Asp	Gly	Val	Val	Gly	Leu	Asp	Asp	
30		225				230					235					240	
	Phe	Arg	Lys	Ser	Gln	Glu	Leu	Arg	Val	Trp	Pro	Gly	Tyr	Asp	Tyr	Val	
					245					250					255		
35	Gly	Trp	Ser	Asn	His	Ser	Phe	Ser	Ser	Gly	Tyr	Val	Glu	Met	Glu	Phe	
				260					265					270			
	Glu	Phe	Asp	Arg	Leu	Arg	Ala	Phe	Gln	Ala	Met	Gln	Val	His	Cys	Asn	
			275					280					285				
40	Asn	Met	His	Thr	Leu	Gly	Ala	Arg	Leu	Pro	Gly	Gly	Val	Glu	Cys	Arg	
		290					295					300					
	Phe	Arg	Arg	Gly	Pro	Ala	Met	Ala	Trp	Glu	Gly	Glu	Pro	Met	Arg	His	
45		305				310					315					320	
	Asn	Leu	Gly	Gly	Asn	Leu	Gly	Asp	Pro	Arg	Ala	Arg	Ala	Val	Ser	Val	
					325					330					335		
50	Pro	Leu	Gly	Gly	Arg	Val	Ala	Arg	Phe	Leu	Gln	Cys	Arg	Phe	Leu	Phe	
				340					345					350			
	Ala	Gly	Pro	Trp	Leu	Leu	Phe	Ser	Glu	Ile	Ser	Phe	Ile	Ser	Asp	Val	
			355					360					365				
55	Val	Asn	Asn	Ser	Ser	Pro	Ala	Leu	Gly	Gly	Thr	Phe	Pro	Pro	Ala	Pro	
		370					375					380					
	Trp	Trp	Pro	Pro	Gly	Pro	Pro	Pro	Thr	Asn	Phe	Ser	Ser	Leu	Glu	Leu	
60		385				390					395					400	
	Glu	Pro	Arg	Gly	Gln	Pro	Arg	Pro	Val	Ala	Lys	Ala	Glu	Gly	Ser	Pro	
					405					410					415		
65	Thr	Ala	Ile	Leu	Ile	Gly	Cys	Leu	Val	Ala	Ile	Ile	Leu	Leu	Leu	Leu	
				420					425					430			
	Leu	Ile	Ile	Ala	Leu	Met	Leu	Trp	Arg	Leu	His	Trp	Arg	Arg	Leu	Leu	
				435				440					445				

	Ser	Lys	Ala	Glu	Arg	Arg	Val	Leu	Glu	Glu	Glu	Leu	Thr	Val	His	Leu
	450						455					460				
5	Ser	Val	Pro	Gly	Asp	Thr	Ile	Leu	Ile	Asn	Asn	Arg	Pro	Gly	Pro	Arg
	465					470					475					480
	Glu	Pro	Pro	Pro	Tyr	Gln	Glu	Pro	Arg	Pro	Arg	Gly	Asn	Pro	Pro	His
					485					490					495	
10	Ser	Ala	Pro	Cys	Val	Pro	Asn	Gly	Ser	Ala	Leu	Leu	Leu	Ser	Asn	Pro
				500					505					510		
	Ala	Tyr	Arg	Leu	Leu	Leu	Ala	Thr	Tyr	Ala	Arg	Pro	Pro	Arg	Gly	Pro
			515					520					525			
15	Gly	Pro	Pro	Thr	Pro	Ala	Trp	Ala	Lys	Pro	Thr	Asn	Thr	Gln	Ala	Tyr
		530					535					540				
20	Ser	Gly	Asp	Tyr	Met	Glu	Pro	Glu	Lys	Pro	Gly	Ala	Pro	Leu	Leu	Pro
	545					550					555					560
	Pro	Pro	Pro	Gln	Asn	Ser	Val	Pro	His	Tyr	Ala	Glu	Ala	Asp	Ile	Val
					565					570					575	
25	Thr	Leu	Gln	Gly	Val	Thr	Gly	Gly	Asn	Thr	Tyr	Ala	Val	Pro	Ala	Leu
				580					585					590		
	Pro	Pro	Gly	Ala	Val	Gly	Asp	Gly	Pro	Pro	Arg	Val	Asp	Phe	Pro	Arg
			595					600					605			
30	Ser	Arg	Leu	Arg	Phe	Lys	Glu	Lys	Leu	Gly	Glu	Gly	Gln	Phe	Gly	Glu
		610					615					620				
35	Val	His	Leu	Cys	Glu	Val	Asp	Ser	Pro	Gln	Asp	Leu	Val	Ser	Leu	Asp
	625					630					635					640
	Phe	Pro	Leu	Asn	Val	Arg	Lys	Gly	His	Pro	Leu	Leu	Val	Ala	Val	Lys
					645					650					655	
40	Ile	Leu	Arg	Pro	Asp	Ala	Thr	Lys	Asn	Ala	Arg	Asn	Asp	Phe	Leu	Lys
				660					665					670		
	Glu	Val	Lys	Ile	Met	Ser	Arg	Leu	Lys	Asp	Pro	Asn	Ile	Ile	Arg	Leu
			675					680					685			
45	Leu	Gly	Val	Cys	Val	Gln	Asp	Asp	Pro	Leu	Cys	Met	Ile	Thr	Asp	Tyr
		690					695					700				
50	Met	Glu	Asn	Gly	Asp	Leu	Asn	Gln	Phe	Leu	Ser	Ala	His	Gln	Leu	Glu
	705					710					715					720
	Asp	Lys	Ala	Ala	Glu	Gly	Ala	Pro	Gly	Asp	Gly	Gln	Ala	Ala	Gln	Gly
					725					730					735	
55	Pro	Thr	Ile	Ser	Tyr	Pro	Met	Leu	Leu	His	Val	Ala	Ala	Gln	Ile	Ala
				740					745					750		
	Ser	Gly	Met	Arg	Tyr	Leu	Ala	Thr	Leu	Asn	Phe	Val	His	Arg	Asp	Leu
			755					760					765			
60	Ala	Thr	Arg	Asn	Cys	Leu	Val	Gly	Glu	Asn	Phe	Thr	Ile	Lys	Ile	Ala
							775					780				
65	Asp	Phe	Gly	Met	Ser	Arg	Asn	Leu	Tyr	Ala	Gly	Asp	Tyr	Tyr	Arg	Val
	785					790					795					800
	Gln	Gly	Arg	Ala	Val	Leu	Pro	Ile	Arg	Trp	Met	Ala	Trp	Glu	Cys	Ile
					805					810					815	

33

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 7 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

Arg Asp Leu Ala Ala Arg Asn
1 5

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 32 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

GGAATTCTAG AMGSGACYTG GCVGCBAGRA AC

(2) INFORMATION FOR SEQ ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 7 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Modified-site
- (B) LOCATION: 5..one-of(2)
- (D) OTHER INFORMATION: /note= "Xaa at position 5 is Tyr or Phe"

(ix) FEATURE:

- (A) NAME/KEY: Modified-site
- (B) LOCATION: 7..one-of(2)
- (D) OTHER INFORMATION: /note= "Xaa at position 7 is Val or Ile"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

Asp Val Trp Ser Xaa Gly Xaa
1 5

(2) INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 32 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

GGAATTCTAG ACACSCCRWA RSWCCASACR TC

(2) INFORMATION FOR SEQ ID NO: 7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 35 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

GACTCGAGTC GACATCGATT TTTTTTTTTT TTTT

(2) INFORMATION FOR SEQ ID NO: 8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

GTTGGGGAAA ATTCACCAT C

(2) INFORMATION FOR SEQ ID NO: 9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

GGCATGAGCC GGAACCTC

(2) INFORMATION FOR SEQ ID NO: 10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

TCTCGATGAC CTGCTCGTCG GTGAGCTGCC

(2) INFORMATION FOR SEQ ID NO: 11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

CAGTCTCACT GCCTCTATTA GAGGTGATGG

36

(2) INFORMATION FOR SEQ ID NO: 12:

(i) SEQUENCE CHARACTERISTICS:

- 5 (A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

AAGCTGAGGA CAAGTACAGG AGCAGGCACA

30

15

(2) INFORMATION FOR SEQ ID NO: 13:

(i) SEQUENCE CHARACTERISTICS:

- 20 (A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

25

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

TCAATGTGGC AGCTACACTA GAGATGTCCA

30

CLAIMS:

1. A method of screening a substance for potential utility as a therapeutic agent in the treatment of cancer which
5 comprises providing a standard system in which a protein tyrosine kinase or an active fragment thereof is able to develop a measurable effect, allowing the protein tyrosine kinase to develop that effect in the presence and absence of the said substance and measuring that effect, ability to
10 produce significant inhibition of the effect being taken as an indication of potential utility as a therapeutic agent, wherein the protein tyrosine kinase is characterised by the amino acid sequence shown in
SEQ ID NO. 1
15 or an amino acid sequence showing a significant degree of homology thereto.
2. A method as claimed in Claim 1 wherein the protein tyrosine kinase or an active fragment thereof is derived from
20 a recombinant expression system or a cellular source in which the protein is overexpressed.
3. A method as claimed in claim 2 wherein the recombinant expression system is an insect baculovirus, an immortalised
25 human breast cell line or a rodent fibroblast cell line.
4. A method as claimed in claim 2 or 3 wherein the measurable effect is protein tyrosine kinase activity which is measured by measuring autophosphorylation or
30 phosphorylation of a model substrate, by following incorporation of radioactive phosphate or by using anti-phosphotyrosine antibodies.
5. A method as claimed in Claim 1 wherein the measurable
35 effect is a phenotypic alteration conferred by expression of the protein tyrosine kinase or an active fragment thereof in a recombinant or non-recombinant cell line.

6. A method as claimed in Claim 1 wherein the measurable effect is the tumorigenic effect of a recombinant or non-recombinant cell line expressing the protein tyrosine kinase or an active fragment thereof when used as a xenograft in an animal.

7. A method as claimed in Claim 5 or 6 wherein the recombinant cell line is a primary immortalised human breast cell line or a rodent fibroblast cell line.

8. A therapeutic agent identified by a method as claimed in any of Claims 1 to 7.

9. A therapeutic agent as claimed in Claim 8 which is a chemical molecule of relatively low molecular weight.

10. A therapeutic agent as claimed in Claim 9 which is a staurosporine analogue, a tyrphostin or a flavenoid.

11. A therapeutic agent as claimed in Claim 8 which is a macromolecule.

12. A therapeutic agent as claimed in Claim 11 which is an antibody.

13. A therapeutic agent as claimed in claim 8 which is a peptide.

14. A therapeutic agent as claimed in claim 8 which is an antisense oligonucleotide capable of blocking translation of the mRNA encoding the protein tyrosine kinase.

15. A method for detecting a tyrosine kinase having an amino acid sequence as defined in

SEQ ID NO. 1

or a fragment thereof which comprises contacting a test sample with a specific antibody raised against an antigen from the

said amino acid sequence and determining whether there is any antigen antibody binding within the test sample.

- 5 16. A method for the detection in a sample of a nucleic acid sequence encoding a tyrosine kinase which comprises subjecting the sample or nucleic acid isolated therefrom to a method capable of detecting a nucleic acid sequence therein identical or substantially homologous to a defined nucleic acid sequence and identifying any sequence identical or substantially
10 homologous to all or part of the DNA sequence defined in SEQ ID NO. 1, or mRNA which would be the transcription product of such DNA.
- 15 17. A method as claimed in Claim 16 wherein DNA is detected by Southern blotting and/or PCR using appropriate primers.
- 20 18. A method as claimed in Claim 16 wherein RNA is detected by Northern blotting, RNase protection studies and direct nucleic acid *in situ* hybridisation.
- 25 19. An enzyme-substrate complex which comprises a protein tyrosine kinase characterised by the amino acid sequence shown in SEQ ID NO. 1 or an amino acid sequence showing a significant degree of
30 homology thereto and a therapeutic agent capable of modulating the activity of said protein tyrosine kinase.
- 35 20. A complex as claimed in Claim 19 wherein the therapeutic agent is a chemical molecule of relatively low molecular weight.
21. A complex as claimed in Claim 20 wherein the therapeutic agent is a staurosporine analogue, a tyrphostin or a
35 flavenoid.
22. A complex as claimed in Claim 19 wherein the therapeutic

agent is a macromolecule.

23. A complex as claimed in Claim 22 wherein the therapeutic agent is an antibody.

5

24. A complex as claimed in Claim 19 wherein the therapeutic agent as a peptide.

25. A complex as claimed in Claim 19 wherein the therapeutic agent is an antisense oligonucleotide capable of blocking translation of the mRNA encoding the protein tyrosine kinase.

10

1/5

FIG. 1

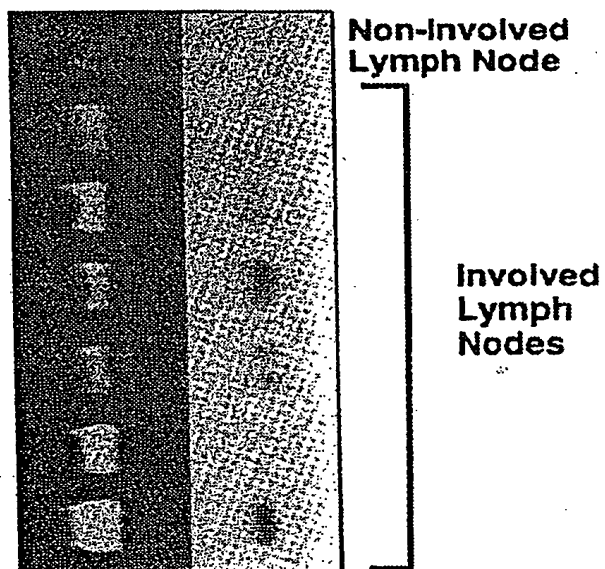
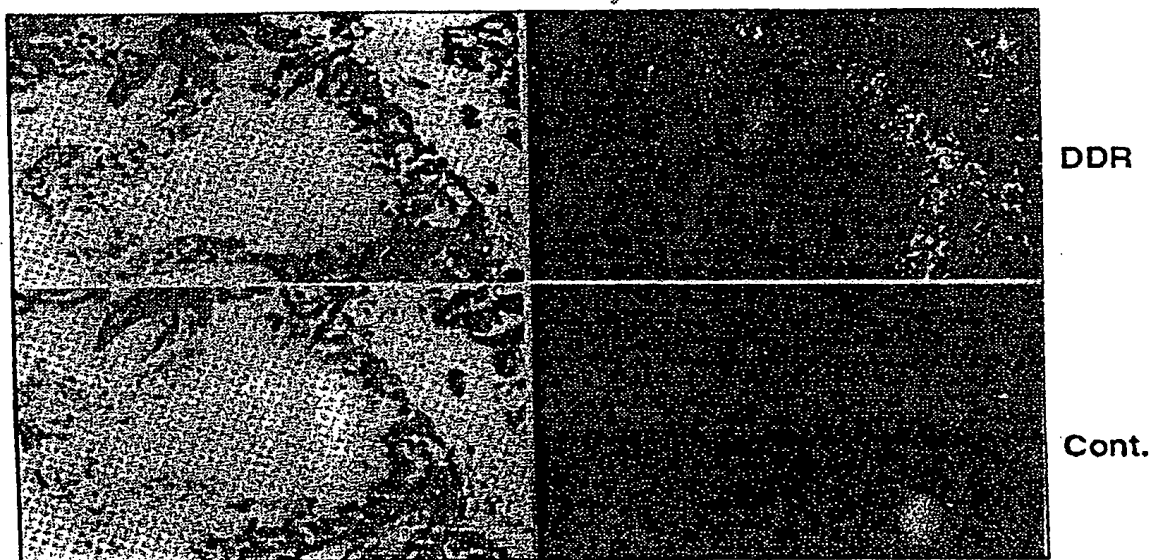


FIG. 2



2/5

FIG. 4

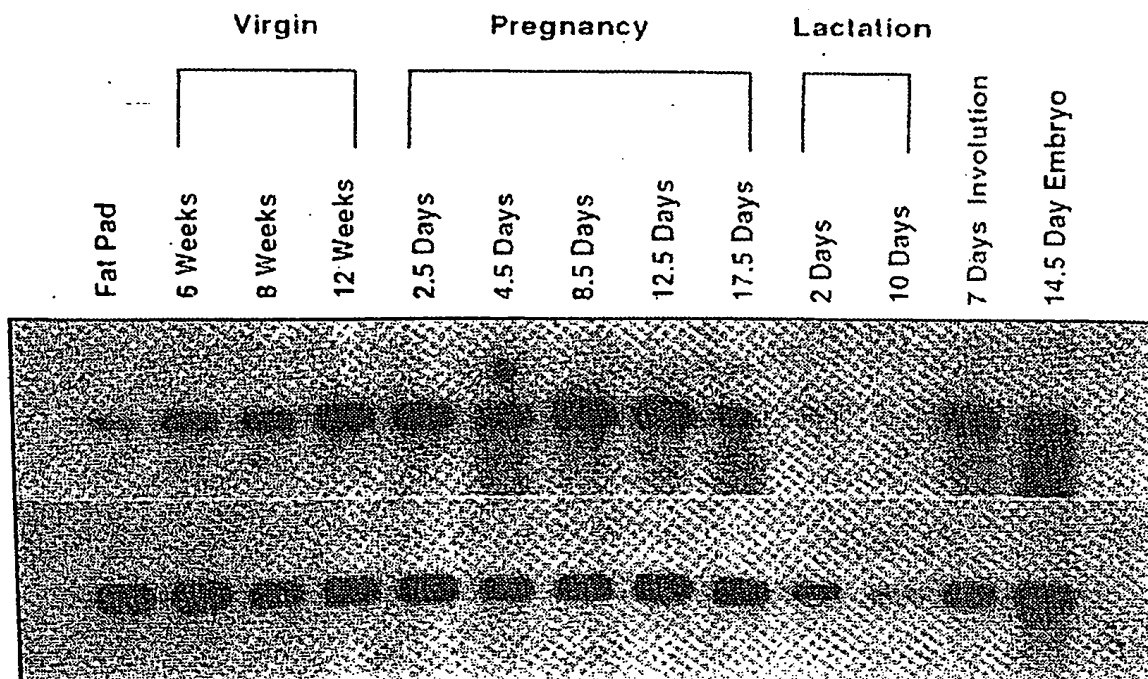
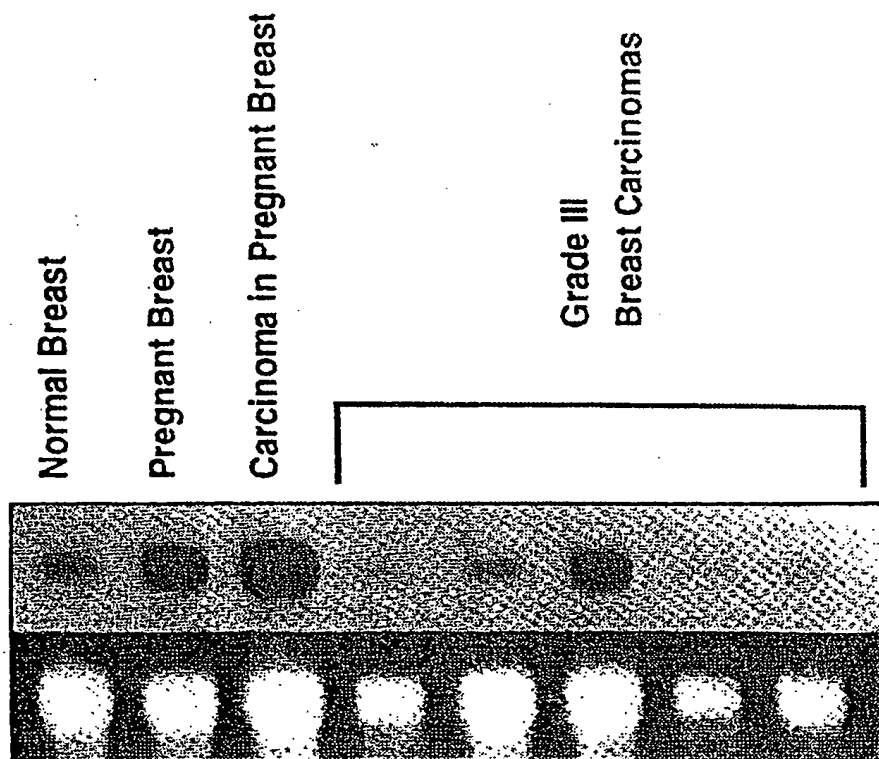


FIG. 3



3/5

FIG. 5

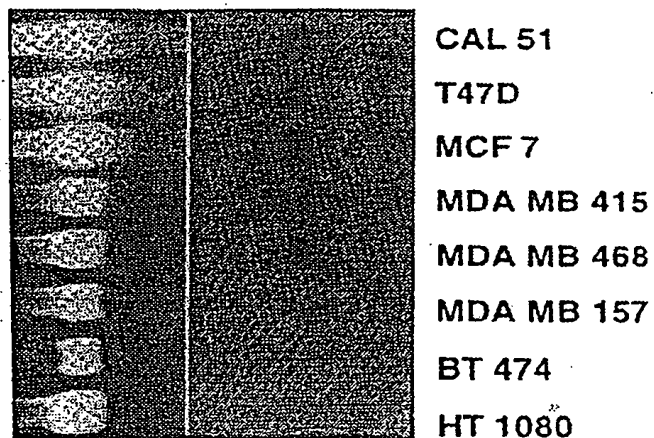
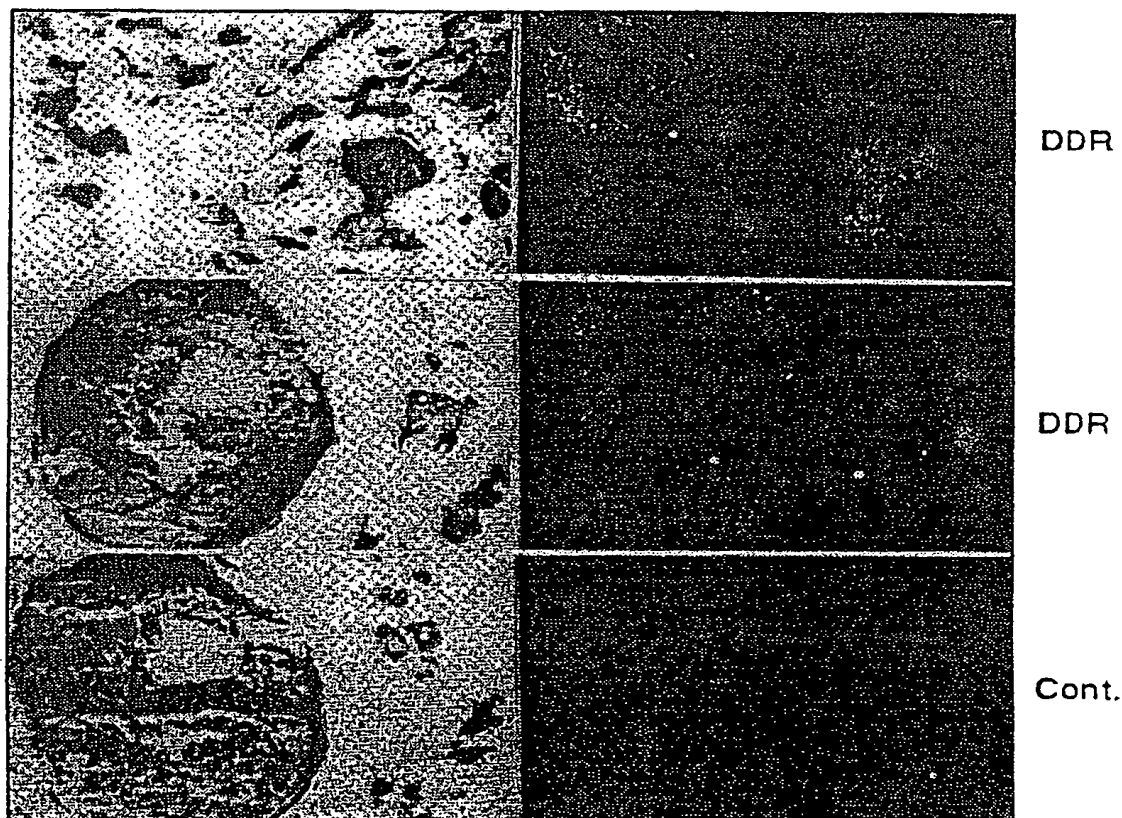
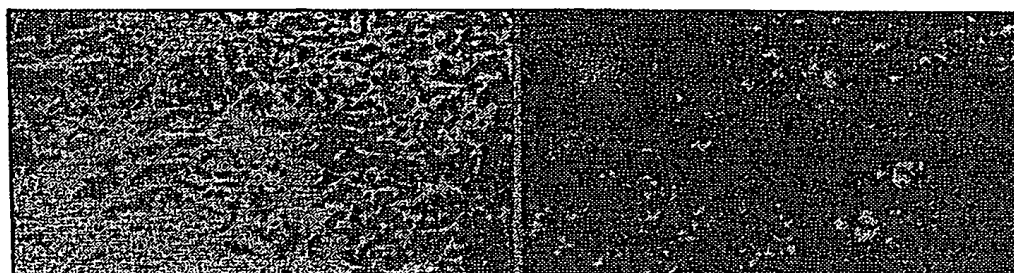


FIG. 6a



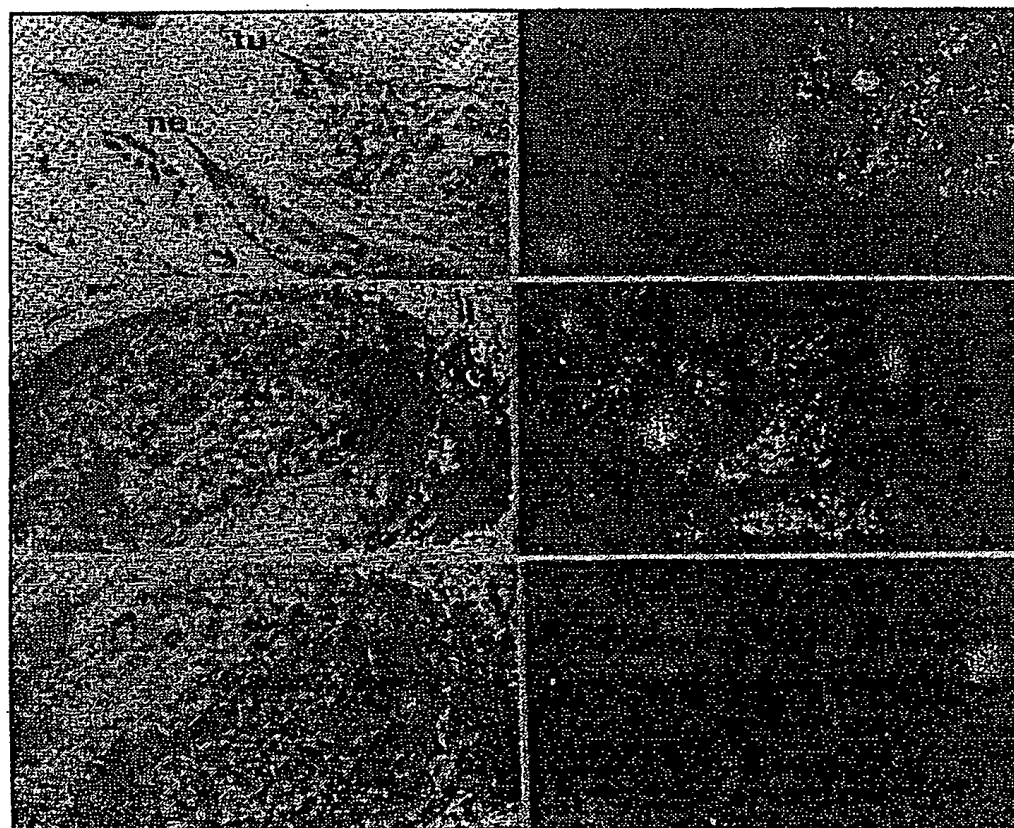
4/5

FIG. 6b



DDR

FIG. 6c



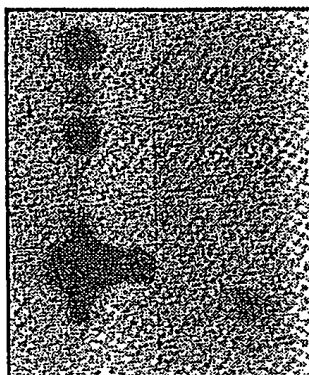
DDR

DDR

Cont.

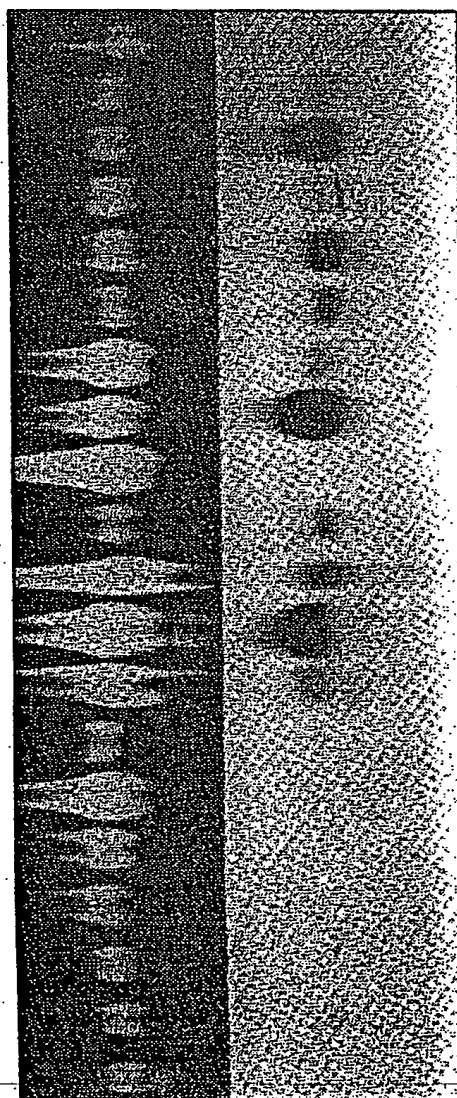
5/5

FIG. 7a



Heart
Brain
Placenta
Lung
Liver
Skeletal Muscle
Kidney
Pancreas

FIG. 7b



Oesophagus
Stomach
Gall Bladder
Ileum
Caecum
Colon
Tonsil
Thyroid
Spleen
Kidney Pelvis
Kidney Medulla
Kidney Cortex
Lung
Skin
Testis
Prostate
Myometrium
Endometrium
Breast
Bone Marrow

SUBSTITUTE SHEET (RULE 26)

INTERNATIONAL SEARCH REPORT

International Application No.
PCT/GB 94/01480

A. CLASSIFICATION OF SUBJECT MATTER
 IPC 6 G01N33/574 C12Q1/48 A61K31/55 A61K31/275 A61K31/35
 A61K39/395 A61K38/00 A61K31/70 G01N33/573 C12Q1/68
 C12N9/12 C12N15/11

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
 IPC 6 G01N C12Q A61K C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF USA, vol.90, June 1993, WASHINGTON US pages 5677 - 5681 JEFFREY D. JOHNSON ET AL. 'A receptor tyrosine kinase found in breast carcinoma cells has an extracellular discoidin I-like domain' cited in the application see the whole document ---	1-25
Y	DRUGS OF THE FUTURE, vol.17, no.2, 1992 pages 119 - 131 TERRENCE R. BURKE, JR. 'Protein-tyrosine kinase inhibitors' cited in the application see the whole document --- -/-	1-25

☒ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

* Special categories of cited documents :

- *A* document defining the general state of the art which is not considered to be of particular relevance
- *E* earlier document but published on or after the international filing date
- *L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- *O* document referring to an oral disclosure, use, exhibition or other means
- *P* document published prior to the international filing date but later than the priority date claimed

- *T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- *X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- *Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- *&* document member of the same patent family

Date of the actual completion of the international search

25 October 1994

Date of mailing of the international search report

0 8. 12 94

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
 NL - 2280 HV Rijswijk
 Tel. (+31-70) 340-2040, Tx. 31.651 epo nl,
 Fax: (+31-70) 340-3016

Authorized officer

Döpfer, K-P

INTERNATIONAL SEARCH REPORT

Intern al Application No

PCT/GB 94/01480

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	JOURNAL OF MEDICINAL CHEMISTRY, vol.34, 1991, WASHINGTON US pages 1896 - 1907 AVIV GAZIT ET AL. 'Tyrphostins. 2. Heterocyclic and alpha-substituted Benzylidenmalononitrile Tyrphostins as Potent Inhibitors of EGF Receptor and ErbB2/neu Tyrosine Kinases' cited in the application see the whole document ---	1-10, 19-21
Y	PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF USA, vol.89, May 1992, WASHINGTON US pages 4285 - 4289 PAUL CARTER ET AL. 'Humanization of an anti-p185(HER2) antibody for human cancer therapy' cited in the application see the whole document ---	1-7, 11-13, 15,19, 22-25
P,A	ONCOGENE, vol.8, no.12, December 1993, BASINGSTOKE GB pages 3433 - 3440 THOMAS KARN ET AL. 'Stzructure, expression and chromosomal mapping of TKT from man and mouse: a new subclass of receptor tyrosine kinases with a factor VIII-like domain' see the whole document ---	1-8, 11-19, 22-25
P,A	ONCOGENE, vol.9, no.1, January 1994, BASINGSTOKE GB pages 211 - 219 JOSE L. PEREZ ET AL. 'Identification and chromosomal mapping of a receptor tyrosine kinase with a putative phospholipid binding sequence in its ectodomain' see the whole document ---	1-8, 11-19, 22-25
P,A	ONCOGENE, vol.8, no.10, October 1993, BASINGSTOKE GB pages 2731 - 2739 MARIELBA ZERLIN ET AL. 'NEP: a novel receptor-like tyrosine kinase expressed in proliferating epithelia' cited in the application see the whole document ---	1-8, 11-19, 22-25
1 A	EP,A,0 504 914 (E.R. SQUIBB & SONS, INC.) 23 September 1992 see the whole document ---	1-8, 11-19, 22-25

	---/---	

INTERNATIONAL SEARCH REPORT

International Application No
PCT/GB 94/01480

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO,A,90 10234 (EURO-DIAGNOSTICS B.V.) 7 September 1990 see the whole document ---	1,4
A	SCIENCE, vol.241, no.4861, 1 July 1988, LANCASTER,PA, US pages 42 - 52 STEVEN H. HANKS, ANNE MARIE QUINN, TONY HUNTER 'The Protein Kinase Family: Conserved Features and Deduced Phylogeny of the Catalytic Domains' cited in the application ---	
A	JOURNAL OF CLINICAL ONCOLOGY, vol.10, no.7, July 1992, NEW YORK, NY, US pages 1049 - 1056 B.A. GUSTERSON ET AL. 'Prognostic Importance of c-erbB-2-Expression in Breast Cancer' cited in the application -----	

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/GB 94/01480

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
EP-A-0504914	23-09-92	JP-A- 5273208	22-10-93
WO-A-9010234	07-09-90	NL-A- 8900529	01-10-90
		AU-A- 5263190	26-09-90
		EP-A- 0461174	18-12-91
		JP-T- 4504171	23-07-92